Optimality conditions for nonsmooth optimization problems via generalised derivatives

Sara Hassani

This thesis is submitted in total fulfilment of the requirement for the degree of Doctoral of Philosophy

Faculty of Science and Technology
Federation University Australia

PO Box 663
University Drive, Mount Helen
Ballarat, VIC 3353, Australia.

Thesis Supervisors:
Dr. Musa Mammadov, A/Prof. Alex Kruger

October 2016

Abstract

The motivation for generalizing the concept of subdifferential is inspired by many real world applications involving nonsmooth functions. These generalizations are very important tools for deriving optimality conditions.

In this thesis, first, we consider optimality conditions for optimization problems by introducing new notions of supporting cones including local and global supporting cones. In the development of these concepts, we use the notion of the augmented normal cone that has been frequently considered in the literature. These introduced concepts of supporting cones together with the so-called weak subdifferentials are used to derive optimality conditions for local and global optima. Then, the obtained results are extended to reflexive Banach spaces.

Next, optimization problems for topical functions are investigated. They arise in a remarkable variety of mathematical disciplines such as matrices over the max-plus semiring, Leontieff substitution systems of mathematical economics, dynamic programming operators of games and of Markov decision processes and nonlinear operators arising in matrix scaling problems and demographic modelling. In this thesis, we present necessary and sufficient conditions for global maxima of the difference of two strictly topical functions defined on a semimodule by using superdifferential of extended valued topical functions.

Finally, we devote the last section to infinite horizon optimization problems, which have many applications in real world problems; for example, in capacity expansion, equipment replacement and production planning. The objective of these problems is to find a sequence of decisions such that the associated cost over an unbounded horizon is optimal. In this

thesis, we investigate the mathematical formulation for the horizon optimization that has been commonly studied in the literature. This involves the description of a mathematical model to describe the sequence of decisions and associated cost as well as the metric in the space of infinite decisions that are used to study the convergence of the minimizing sequence. Moreover, we investigate optimality conditions for the minimum cost of the infinite horizon optimization by introducing two notions, contingent cone and upper contingent derivative.

Statement of Authorship

Except where explicit reference is made in the text of the thesis, this thesis contains no material published elsewhere or extracted in whole or in part from a thesis by which I have qualified for or been awarded another degree or diploma. No other persons work has been relied upon or used without due acknowledgement in the main text and bibliography of the thesis.

Sara Hassani

October 2016

Acknowledgement

I would like to express my special appreciation and thanks to my principal supervisor, Dr Musa Mammadov, for his strong support, kind consideration, helpful guidance and valuable time throughout my study. His scientific expertise, continuous encouragement and support of my research have been essential for my success. I also thank my associate supervisor, Assoc. Prof. Alex Kruger for his encouragement and help throughout my PhD study.

Special thank to my dear Ehsan for his lovely support and inspiration. I know without his support I could not achieve my ambitions.

Dedication

To my Ehsan, as always

Contents

Chapter 1

Introduction

Constrained optimization problems occur naturally in many applications. A simplified form of such problems can be formulated as finding an optimal value of function f over Ω, that is,

$$\text{Minimize/Maximize} \quad f(\mathbf{x}), \qquad \text{subject to} \qquad \mathbf{x} \in \Omega, \tag{1.1}$$

where Ω is referred to a feasible set. Optimization problems can be divided into large groups such as convex and nonconvex problems or smooth and nonsmooth problems. Nonconvex and nonsmooth optimization problems, which are studied in this thesis, generalize some of the results from [59].

The main results obtained throughout this thesis deal with notions of subdifferentials, superdifferentials and normal cones, which have a crucial role in nonsmooth optimization. The concept of subdifferential of convex functions, which was introduced by Moreau and Rockafellar [93], is a generalization of the derivative to functions that are not differentiable. Moreover, the generalized subdifferential for nonconvex functions was introduced by Clarke and Rockafellar [21, 22, 25, 27]. Constructing subdifferentials in the standard way and in terms of directional derivatives gives convex subdifferentials and normal cones. However, convexity is sometimes a very restrictive assumption, so there is a clear need to study subdifferentials and normal cones that are not necessarily convex. In this direction, Kruger and Mordukhovich

[62, 63, 64, 67, 68, 82] introduced the limiting subdifferential that, in general, is not convex. These generalizations of the derivative and normal cone have been used in many applications including optimal control, mathematical programming and constrained optimization.

Demyanov and Rubinov ([30]), by using subdifferential and superdifferential, introduced the concept of quasidifferentiability. It is a convenient tool for constructing a linearization of the directional derivative for a broad class of nonsmooth functions. One of the main ideas related to quasidifferentiability is Abstract Convexity, which opens the way for extending some main ideas and results from classical convex analysis to much more general classes of functions, mappings and sets.

In the classical convex analysis, the elements of subdifferentials are continuous linear functions, but in the case of abstract convexity, the elements of abstract subdifferentials are not necessarily linear functions. For example, a special class of concave functions was introduced by Azimov and Gasimov [2, 3], and by using nonlinear functions, the notion of weak subdifferential was introduced.

The concept of normal cones to convex sets and their properties have been investigated in [76, 83, 93]. Generalizations of normal cones to nonconvex sets were introduced in [83]. There are approaches to define normal cones based on subdifferentials. For instance, normal cone can be represented by subdifferential of the distance function.

Normal cones and subdifferentials play a significant role in dealing with optimality conditions in optimization problems. As an example, a well-known optimality condition for a local minimizer of a real valued convex function f at a point $\bar{x}$ on a convex set Ω is $0 \in N(\bar{x}; \Omega) + \partial f(\bar{x})$.

In this thesis, two important classes of optimization problems containing minimizing of the difference of two topical functions and minimizing of the cost of infinite horizon are presented.

Topical functions have many applications in various parts of applied mathematics, in particular, in the modeling of discrete event systems [42, 43]. They are also interesting as a tool in the study of the so-called downward sets. Downward sets arise in the study of some problems of mathematical economics and game theory and also in the study of inequality systems involving increasing functions [96]. Recently, topical functions $f : X \to K$ and related classes of functions have been studied in [20, 80, 96, 105], where X is a b-complete idempotent semimodule over a b-complete idempotent semifield K. Extended valued topical functions with values in $\bar{K} := K \cup \tau$, where $\tau := \sup K$ (possibly does not belong to K), have been investigated in [6, 108, 109]. In this thesis, one of our major goals is characterization of the superdifferential of extended valued topical functions defined on a semimodule with values in a semifield, and necessary and sufficient conditions for global optimality of the difference of two strictly topical functions defined on a semimodule.

Many important planning problems such as capacity expansion, equipment replacement and production planning involve sequences of related decisions over an infinite time horizon. The mathematical formulation of such problems lead to infinite horizon optimization, which is the optimization problem of selecting a sequence of decisions such that the associated cost over an unbounded horizon is optimal [9, 10]. Since the cost over an unbounded horizon may be infinite or diverge, different optimality criteria apart from minimal total cost are required such as discounting factor, average cost and overtaking optimality.

In many studies an optimal solution/trajectory to an infinite horizon problem is approximated by a sequence of finite horizon optimal solutions. Many algorithms are developed for finding optimal solution and the results are applied for instance to undiscounted Markov decision processes and scheduling production problems [10, 101]. In this thesis, we study stability of trajectories in infinite horizon optimization. We also introduce the notions of contingent cone and upper contingent derivative to derive optimality conditions for this class of optimization problems.

1.1 Motivation for research

Generalized normal cones and subdifferentails have many applications in mathematical programming and optimal control. There are a large number of monographs and articles investigating and constructing various generalizations of normal cones and subdifferentials (for details, see Chapter 2). The main purpose of these generalizations is to establish optimality conditions for a broad class of nonsmooth optimization problems. One of the main goals of this thesis is to study necessary and sufficient optimality conditions for a broad class of nonconvex optimization problems. A new supporting cone based on augmented normal cones is introduced and by using weak subdifferentials, optimality conditions for a large class of nonconvex problems are given.

Characterizations of superdifferential of topical functions as well as optimality conditions for minimizing of topical functions defined on semimodules have not been studied, yet. Therefore, we first give characterizations of the superdifferential of extended valued topical functions defined on a semimodule with values in a semifield. Next, as an application, we present necessary and sufficient conditions for global maximizers of the difference of two strictly topical functions defined on a semimodule.

Infinite horizon optimization problems are used in many applications including planning problems, capacity expansion and equipment replacement. Many algorithms were developed for finding optimal solution of such problems. Therefore, the stability of optimal trajectories is the main concern when investigating the convergence of sequences to optimal trajectory in algorithms. In this thesis, we study the classical and ideal convergence of a minimizing sequence of trajectories as well as the optimality condition for infinite horizon optimization.

1.2 Outline of thesis

This thesis focuses on generalizing augmented normal cones to establish optimality conditions for a broad class of nonconvex nonsmooth optimization problems. Based on augmented normal cones, new notions of supporting cones are introduced for local and global supporting cones. The new notions are employed to define a property of sets at a point, which called conic gap, and then to characterize nonconvex sets with respect this property.

Next, we establish the necessary and sufficient conditions considered in [59] for local optima, for a broad class of finite dimensional normed spaces in terms of introduced notions and weak subdifferentials. Similar investigations are done to derive optimality conditions for global optima. Finally the obtained results for optimality conditions of global solutions are extended in reflexive Banach spaces. In the sequel, we investigate optimality conditions for a nonconvex class of optimization problems, minimizing difference of two topical functions, by using superdifferentials.

The other optimization problem that is considered, is minimizing of the cost over infinite horizon in infinite horizon optimization problems. Since stability of optimal trajectories is an important concern when studying infinite horizon optimization problems, first, we establish the convergence of a sequence of minimizing trajectories in the sense of ideals and their particular case called statistical convergence. Then, based on two concepts of contingent cone and upper contingent cone derivative, optimality conditions for infinite horizon optimization, are obtained.

1.3 Structure of thesis

In this section, a brief description of the format of the presented thesis (PhD by publication) is given. An introduction is followed by nine papers with different status of publications, one accepted paper, seven published and one submitted papers. In chapter two, a literature review of generalizations of differentials and normal cones as well as some applications are provided. Necessary and sufficient conditions via weak subdifferentials and augmented normal cone, for local and global optimality, is presented in chapter three and four respectively. The optimality conditions of two important optimization problems are considered in chapter five and six.

The list of nine papers is presented in the next section after the flow chart.

Chapter 3: Necessary and sufficient conditions for local optimality via weak subdifferentials

Section 3.1: Paper 1

Title: Relation between weak subdifferentials and Hadamard lower directional derivatives
Status of the paper: Published
Outcome: The relation between weak subdierentials and functions which have Hadamard directional derivative is investigated.

Section 3.2: Paper 2

Title: Some geometrical properties of non-convex sets in finite dimensions
Status: Published
Outcome: A new local supporting function is introduced and geometrical properties of non-convex sets are considered in this paper.

Section 3.3: Paper 3

Title: Necessary and sufficient conditions for local optimality via weak subdifferentials.
Status of the paper: Published
Outcome: In this paper, necessary and sufficient conditions for local optimality in finite dimensional normed spaces in terms of weak subdifferential are studied.

Chapter 4: Necessary and sufficient conditions for global optimality via sigma supporting cone

Section 4.1: Paper 4

Title: Characterizing non-convex sets with conic gap via sigma supporting cone
Status of the paper: Published
Outcome: In this paper, a new global supporting function is introduced in Banach spaces.

Section 4.2: Paper 5

Title: Sigma supporting cone and optimality conditions in non-convex problems
Status of the paper: Published
Outcome: The paper is presented the optimality conditions for a broader class of finite dimensional normed spaces in terms of weak subdifferentials.

Section 4.3: Paper 6

Title: Optimality condition via weak subdierentials in reflexive Banach spaces
Status of the paper: Published
Outcome: The paper is presented the optimality conditions in reflexive Banach spaces.

Chapter 5: Optimality conditions for difference of topical functions and infinite horizon optimization

Section 5.1: Paper 7

Title: Characterizations of minimal elements of topical functions on semimodules with applications
Status of the paper: Published
Outcome: Optimality conditions for difference of two topical functions defined on semimodules, by using superdifferentials are considered.

Section 5.2: Paper 8

Title: Convergence of trajectories in infinite horizon optimization
Status of the paper: Accepted
Outcome: The paper is devoted to the convergence of a sequence of minimizing trajectories in infinite horizon optimization problems.

Section 5.3: Paper 9

Title: Optimality conditions in infinite horizon optimization by contingent derivative
Status of the paper: Submitted
Outcome: Two new concepts are introduced in this section. Then, the optimality conditions are derived using new notions.

1.4 List of papers

[1] S. Hassani and M. Mammadov (2013) Relation between weak subdifferentials and Hadamard lower directional derivatives, The 44th Annual Iranian Mathematics Conference, 348-351.

[2] S. Hassani and M. Mammadov (2014) Some geometrical properties of non-convex sets in finite dimensions, The 45th Annual Iranian Mathematics Conference, 112-114.

[3] S. Hassani and M. Mammadov (2014) Necessary and sufficient conditions for local optimality via weak subdifferentials, Advances in Theoretical and Applied Mathematics, Volume 9, Number 2, 143-154.

[4] S. Hassani, M. Mammadov and M. Jamshidi (2014) Characterizing non-convex sets with conic gap via sigma supporting cone, Proceedings of the 10th IMT-GT ICMSA, 143-148.

[5] S. Hassani and M. Mammadov (2014) Sigma supporting cone and optimality conditions in non-convex problems, Far East Journal of Mathematical Sciences, Volume 91, Number 2, 169-190.

[6] S. Hassani, M. Mammadov and M. Jamshidi (2017) Optimality condition via weak subdierentials in reflexive Banach spaces, Turkish Journal of Mathematics, 41, 1-8.

[7] S. Hassani and H. Mohebi (2017), Characterizations of minimal elements of topical functions on semimodules with applications, Linear Algebra and its Applications, 520, 104-124.

[8] S. Hassani and M. Mammadov (2016), Convergence of trajectories in infinite horizon optimization, Accepted in International journal of nonlinear analysis and its applications.

[9] S. Hassani and M. Mammadov (2016), Optimality conditions in infinite horizon optimization by contingent derivative, Submitted.

Chapter 2

Literature Review

In this chapter, we aim to make a sketch of what has been brought out in the literature with regard to generalized gradient and normal cone and their characterizations and applications. The last part of this chapter is devoted to the important class of optimization problems called infinite horizon optimization and topical functions. The notations and symbols used throughout this chapter is presented in the next section. Differentiablity properties of convex functions, normal cone to convex sets and subdifferentials of convex functions are discussed in subsections 2.2.1, 2.2.2 and 2.2.3 respectively. In the subsections, 2.3.1, 2.3.2, 2.3.3, 2.3.4 and 2.3.6, a variety of subdifferentials and normal cones and related optimality conditions for nonconvex and nonsmooth optimization problems are considered. Finally, in sections 2.4 and 2.5, we study two important classes of optimization problems where the objective functions are topical functions and cost over infinite horizon optimization problems respectively.

2.1 Notations

Unless otherwise stated, in this section, X is a real Banach space, its dual space is denoted by X^*, while $\langle .,. \rangle$ denotes the pairing between X and X^*. A function $\varphi : X \to [-\infty, +\infty]$ is called convex, if for all $x, y \in X$, $\lambda \in [0,1]$ and $\alpha, \beta \in R$ with $\varphi(x) < \alpha$, $\varphi(y) < \beta$, then $\varphi(\lambda x + (1-\lambda)y) < \alpha\lambda + (1-\lambda)\beta$.

A function $\varphi : X \to [-\infty, +\infty]$ is said to be lower semi-continuous at $a \in X$ if

$$\liminf_{x \to a} \varphi(x) \geq \varphi(a).$$

A function φ is locally Lipschitz at x if

$$|\varphi(u) - \varphi(v)| \leq K\|u - v\|,$$

for some $K > 0$ and for all u, v in some neighborhood of x.

A function $\varphi : X \to (-\infty, +\infty]$ is positively homogeneous if

$$\varphi(\lambda x) = \lambda \varphi(x) \ \ \forall x \in X, \forall \lambda > 0.$$

For a set Ω in X, i_Ω denotes the indicator function of Ω

$$i_\Omega(x) := \begin{cases} 0, & \text{if } x \in \Omega, \\ \infty, & x \notin \Omega. \end{cases}$$

We denote the domain of function $\varphi : X \to [-\infty, +\infty]$ by domφ, which is defined as dom$\varphi := \{x \in X : \varphi(x) < +\infty\}$. The relative interior riΩ of $\Omega \subset X$ is the interior of Ω with respect to affine hull of Ω.

A function $\varphi : X \to (-\infty, +\infty]$ is called proper if dom$\varphi \neq \emptyset$.

2.2 Convex problems

Convex analysis starts with the fact that every closed convex set is Chebyshev set and that if the converse is true [104]. Convex sets and convex functions play an important role in many optimization problems. In this section, various derivatives relative to convex functions are presented. For nondifferentiable functions the notion of subdifferential, that is, a generalization of the ordinary derivative is introduced. Finally, by applying subdifferentials and normal cones, the optimality conditions for convex optimization problems are given.

2.2.1 Differentiablity properties of convex functions

In this subsection, various derivatives relative to extended real-valued convex functions is considered. As the starting point, the case of one variable functions is discussed. Right and left derivatives of a function $\varphi : R \to (-\infty, +\infty]$ are defined by

$$\varphi'_+(\bar{x}) := \lim_{x \to \bar{x}^+} \tfrac{\varphi(x)-\varphi(\bar{x})}{x-\bar{x}},\ \varphi'_-(\bar{x}) := \lim_{x \to \bar{x}^-} \tfrac{\varphi(x)-\varphi(\bar{x})}{x-\bar{x}},$$

where $x \to \bar{x}^+$ and $x \to \bar{x}^-$ mean x approaches to $\bar{x}$ from right and left respectively.

The following result shows right and left derivatives of convex functions [86, Lemma 3.15],[110, Theorem 1.6], [117, Theorem 2.1.5].

Theorem 2.2.1. *Let $\varphi : R \to (-\infty, +\infty]$ be convex. Then φ has a right derivative and a left derivative at every point $\bar{x} \in$ int (dom φ), and*

$$\varphi'_-(\bar{x}) \le \varphi'_+(\bar{x}).$$

Moreover, we have

$$\tau \in [\varphi'_-(\bar{x}), \varphi'_+(\bar{x})] \Longleftrightarrow \tau(x-\bar{x}) \le \varphi(x) - \varphi(\bar{x}).$$

In addition, the set of points at which the convex function $\varphi : R \to R$ is not differentiable is at most countable [86, Proposition 3.16] and [110, Section 1.8].

The classical directional derivative defined by Dini [32] is given by

$$\varphi'(\bar{x}; x) := \lim_{t \downarrow 0} \frac{\varphi(\bar{x}+tx) - \varphi(\bar{x})}{t}. \tag{2.1}$$

Every convex function $\varphi : X \to (-\infty, +\infty]$ admits classical directional derivative in all directions $x \in X$ at any point of $\bar{x}$ in its domain [5, 51, 117].

Another differentiable properties of convex functions are Fréchet and Gâteaux differentiability in general spaces.

Definition 2.2.1. *A function $\varphi : X \to (-\infty, +\infty]$ is said to be Fréchet differentiable (or differentiable) at $\bar{x}$ if there exists $x^* \in X^*$ such that*

$$\lim_{x \to \bar{x}} \frac{\varphi(x) - \varphi(\bar{x}) - \langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} = 0. \tag{2.2}$$

φ is said to be Gâteaux differentiable at $\bar{x}$ if there exists $x^* \in X^*$ such that

$$\lim_{t \downarrow 0} \frac{\varphi(\bar{x} + tx) - \varphi(\bar{x})}{t} = \langle x^*, x \rangle, \quad \forall x \in X. \tag{2.3}$$

Indeed, when the directional derivative $\varphi^{'}(\bar{x}; \cdot)$ is linear, we say φ is Gâteaux differentiable at $\bar{x}$. Gâteaux derivative is denoted by $\nabla\varphi(\bar{x})$.

Fréchet differentiablity implies Gâteaux differentiablility [51, Proposition 1] but the converse is not true. However for convex functions defined on finite dimensional normed vector spaces, Gâteaux and Fréchet differentiability coincide [117]. In the following result [117, Proposition 3.3.6], a necessary and sufficient condition for coinciding Fréchet and Gâteaux derivative of a convex function is given.

Proposition 2.2.2. *Let the proper convex function $\varphi : X \to (-\infty, +\infty]$ be continuous and Gâteaux differentiable on a neighborhood of $\bar{x} \in$domφ. Then φ is Fréchet differentiable at $\bar{x}$ if and only if $\nabla\varphi$ is continuous at $\bar{x}$, where $\nabla\varphi$ is the Gâteaux derivative of φ at $\bar{x}$.*

Several generalizations of directional derivatives are considered in the literature but in some of them when the function is convex they reduce to the classical directional derivative. For example upper Dini directional derivative of φ at $\bar{x} \in$dom φ in direction x defined by

$$\bar{D}\varphi(\bar{x}; x) := \limsup_{t \downarrow 0} \frac{\varphi(\bar{x} + tx) - \varphi(\bar{x})}{t} \tag{2.4}$$

is equal to $\varphi^{'}(\bar{x}; x)$ when φ is convex [73].

2.2.2 Normal cone to convex sets

A convex cone which arises frequently in optimization is the normal cone to a convex set Ω at a point $\bar{x} \in \Omega$, denoted by $N(\bar{x}; \Omega)$, goes back to Minkowski [76]. This is the convex cone of normal vectors, that is vectors d in X^* such that $\langle d, x - \bar{x} \rangle \leq 0$ for all x in Ω. The latter notion was treated by Fenchel in [37] as consisting of the outward normals to the supporting half-spaces to Ω at $\bar{x}$.

Finding the maximum or minimum of a function φ relative to a set Ω are fundamental questions in optimization problems. In this case, the main question may arise is under what conditions, existance of minimum of the function $\varphi : \Omega \to (-\infty, +\infty]$ on the set Ω will be guaranteed. Therefore, the derivative, its generalizations and normal cones are applied to derive optimality conditions.

In the following proposition, an optimality condition for local minimizer of a function φ with respect to the classical directional derivative, Gâteaux derivative and normal cone [11, 14] is presented.

Proposition 2.2.3. (First order necessary condition) *Suppose that Ω is a convex set in $\mathbf{X}$, and that the point $\bar{x}$ is a local minimizer of the convex function $\varphi : \Omega \to (-\infty, +\infty]$. Then for any point $x \in \Omega$, $\varphi^{'}(\bar{x}; x - \bar{x}) \geq 0$. In particular, if φ is Gâteaux differentiable at $\bar{x}$ then the condition $-\nabla\varphi(\bar{x}) \in N(\bar{x}; \Omega)$ holds.*

2.2.3 Subdifferential of convex functions

Since even convex continuous functions may not be differentiable, an important issue is to generalize differentiability. The concept of differentiability is developed in convex analysis to subgradient of a convex function[18, 77, 78, 79, 92].

Definition 2.2.2. *A vector $x^* \in X^*$ is said to be a subgradient of a convex function φ at a point $\bar{x} \in dom\varphi$ if*

$$\varphi(x) \geq \varphi(\bar{x}) + \langle x^*, x - \bar{x} \rangle \ \ \forall x \in X. \tag{2.5}$$

The equation (2.5) which is referred to as the subgradient inequality has a simple geometric meaning when φ is finite at $\bar{x}$: it says that the graph of the affine function $h(x) = \varphi(\bar{x}) + \langle x^*, x - \bar{x} \rangle$ is a non-vertical supporting hyperplane to the convex set epiφ at the point $(\bar{x}, \varphi(\bar{x}))$ i.e. $(x^*, -1) \in N((\bar{x}, \varphi(\bar{x})); \text{epi}\varphi)$. The set of subgradients of φ at $\bar{x}$ is denoted by $\partial\varphi(\bar{x})$ and called (Moreau-Rockafellar) subdifferential. If $\bar{x} \notin dom\varphi$, define $\partial\varphi(\bar{x}) := \emptyset$. If $\partial\varphi(\bar{x})$ is not empty, φ is said to be subdifferentiable at $\bar{x}$. It is clear from definition 2.2.2 that $\partial\varphi(\bar{x})$ is always a weak* closed convex subset of X^* [18].

By Theorem 2.2.1, subdifferential of a convex function $\varphi : R \to (-\infty, +\infty]$ at $\bar{x} \in$int(dom φ) is determined by:

$$\partial\varphi(\bar{x}) = [\varphi'_-(\bar{x}), \varphi'_+(\bar{x})].$$

If φ actually has a unique gradient $x^* = \nabla\varphi(\bar{x})$ at $\bar{x}$ in the sense of Gâteaux (or Fréchet), one would in particular have $\partial\varphi(\bar{x}) = \{\nabla\varphi(\bar{x})\}$ [117, Corollary 2.4.10]. The next result concerns subdifferentiable of functions which are *Gâteaux* differentiable [86, Corollary 3.26].

Proposition 2.2.4. *Let $\varphi : X \to (-\infty, +\infty]$ be a convex function. If φ is finite and continuous at $\bar{x} \in X$ and $\partial\varphi(\bar{x})$ is a singleton x^*, then φ is Gâteaux differentiable at $\bar{x}$ and $x^* = \nabla\varphi(\bar{x})$.*

The next two propositions present conditions under which convex functions are subdifferentiable [5, Proposition 2.36], [86, Corollary 3.28].

Proposition 2.2.5. *If a convex function $\varphi : X \to (-\infty, +\infty]$ is (finite and) continuous at $\bar{x}$, then φ is subdifferentiable at $\bar{x}$, that is, $\partial\varphi(\bar{x}) \neq \emptyset$.*

Proposition 2.2.6. *Let φ be a convex function on a finite-dimensional normed space X and let $\bar{x}$ $\in$ri(dom φ). Then $\partial\varphi(\bar{x})$ is nonempty.*

In the following result, the relation between classical directional derivative and subgradient is given [5, Proposition 2.39].

Proposition 2.2.7. *Let $\varphi : X \to (-\infty, +\infty]$ be a proper convex function. If φ is finite and continuous at $\bar{x}$, then*

$$\varphi^{'}(\bar{x}; x) = \max\{\langle x^*, x\rangle : \ x^* \in \partial\varphi(\bar{x})\},$$

and, in general, one has

$$\partial\varphi(\bar{x}) = \{x^* \in X^* : \ \langle x^*, x\rangle \leq \varphi^{'}(\bar{x}; x), \ \ \forall x \in X\}.$$

In the general theory of convex optimization, the following trivial consequence of Definition 2.2.2 plays an important role. Indeed, it gives a global optimality condition for convex functions by applying subdiffferentials.

Proposition 2.2.8. *If φ is a proper convex function on X, then the minimum (global) of φ over X is attained at the point $\bar{x} \in X$ if and only if $0 \in \partial\varphi(\bar{x})$.*

The next result which is known as Pshenichnyi-Rockafellar theorem concerns an optimality condition of the convex function $\varphi : \Omega \to (-\infty, +\infty]$ which is a consequence of Moreau-Rockafellar theorem.

Theorem 2.2.9. *Let $\varphi : \Omega \to (-\infty, +\infty]$ be a proper convex function and Ω be a convex set. Suppose that either dom$\varphi\cap$int$\Omega \neq \emptyset$, or there exists x_0 $\in$dom$\varphi \cap \Omega$, where φ is continuous at x_0. Then, $\bar{x}$ is a minimum of φ if and only if $\partial\varphi(\bar{x}) \cap (-N(\bar{x}; \Omega)) \neq \emptyset$.*

Let us describe the notion of the normal cone to a convex set in terms of subdifferentials. A simple interpretation of the subdifferential of a function can be given in terms of the normal cone to its epigraph. An important special case of subgradients is the case where φ is the indicator of a non-empty convex set Ω. Then, $x^* \in \partial i_{\Omega}(\bar{x})$ if and only if

$$\langle x^*, x - \bar{x} \rangle \leq 0, \ \ \forall x \in \Omega,$$

i.e. x^* is normal to Ω. Thus $\partial i_\Omega(\bar{x})$ is the normal cone to Ω at $\bar{x}$, i.e. $\partial i_\Omega(\bar{x}) = N(\bar{x}; \Omega)$.

The concept of subgradient extends to ϵ-subgradient which is called approximate subgradient [18]. In practice for example for solving numerically problems using computers it is possible to determine the subgradients only approximately. In this sense the following notion of subgradient reveals itself to be useful.

Definition 2.2.3. *Let φ be a proper convex function. For each $\epsilon > 0$, we define a set $\partial_\epsilon \varphi(\bar{x})$ of "ϵ-subgradients" of φ at $\bar{x}$ by*

$$\partial_\epsilon \varphi(\bar{x}) := \{x^* \in X^* : \ \varphi(x) \geq \varphi(\bar{x}) + \langle x^*, x - \bar{x} \rangle - \epsilon\}. \tag{2.6}$$

Similar to what we have in Proposition 2.2.7, the relation of ϵ-subgradient and ϵ-directional derivative is considered in R^n by Rockafellar [93] and then in general cases by Hiriart-Urruty [48] (see [117, Theorem 2.4.11]).

Theorem 2.2.10. *Let $\varphi : X \to (-\infty, +\infty]$ be a proper convex function, $\bar{x} \in dom\varphi$, $\epsilon \in (0, +\infty)$. Then for any $x \in X$*

$$\varphi'_\epsilon(\bar{x}; x) = \sup\{\langle x^*, x \rangle : \ x^* \in \partial_\epsilon \varphi(\bar{x})\},$$

where

$$\varphi'_\epsilon(\bar{x}; x) = \inf_{t>0} \tfrac{\varphi(\bar{x}+tx) - \varphi(\bar{x}) + \epsilon}{t}.$$

$\varphi'_\epsilon(\bar{x}; x)$ is an extension of the classical directional derivative. We say that $\bar{x} \in \Omega$ is a local ϵ- (minimum) solution of $\varphi : \Omega \to \bar{R}$ if $\varphi(\bar{x}) \leq \inf_{x \in \Omega} \varphi(x) + \epsilon$. We can have optimality conditions similar to Proposition 2.2.8 for local ϵ-(minimum) solution (see [117, Proposition 2.5.9]).

2.3 Nonsmooth and nonconvex problems

Nonsmooth and nonconvex analysis had its origins in the early 1970's when control theorists and nonlinear programmers attempted to deal with necessary optimality conditions for problems with nonsmooth data or nonsmooth functions. To deal with nonsmoothness and nonconvexity, various kinds of generalized derivatives and normal cones have been introduced.

First we present the concept of subdifferentials that encompasses both the notion of derivative and the notion of subdifferential. The main advantages of these concepts are their close relationships with corresponding notions of derivatives and the fact that they provide rather accurate approximations. The following definition introducing Fréchet subdifferential is obtained as a simple one-sided modification of the concept of Fréchet derivative (see Definition 2.2.1).

2.3.1 Fréchet subdifferentials and normal cones

Fréchet subdifferentials have been known for more than forty years. First they were introduced in finite dimensions in [7] (under the name "lower semidifferentials"). Some of their properties in an infinite-dimensional setting is investigated in [60, 61, 65].

Definition 2.3.1. *Given a function $\varphi : X \to \bar{R}$ and finite at $\bar{x} \in X$. The Fréchet subdifferential of φ at $\bar{x}$ is the set of $\partial_F \varphi(\bar{x})$ of $x^* \in X^*$ satisfying the following property: for every $\epsilon > 0$ there exists some $\delta > 0$ such that*

$$\varphi(x) - \varphi(\bar{x}) - \langle x^*, x - \bar{x} \rangle \geq -\epsilon \|x - \bar{x}\|, \ \ \forall x \in B(\bar{x}, \delta). \tag{2.7}$$

In other words, $x^* \in \partial_F \varphi(\bar{x})$ if and only if

$$\liminf_{x \to \bar{x}} \frac{\varphi(x) - \varphi(\bar{x}) - \langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \geq 0. \tag{2.8}$$

If $\partial_F \varphi(\bar{x}) \neq \emptyset$, we say φ is Fréchet subdifferentiable at $\bar{x}$. Define $\partial_F \varphi(\bar{x}) = \emptyset$, if φ is not finite at $\bar{x}$. Fréchet subdifferential is weak* closed and convex and its elements are called Fréchet subgradients (regular subgradients [94]). Obviously, the subdifferential in Definition 2.2.2 is contained in Fréchet subdifferential. If φ is convex, then two subdifferentials coincide. Clearly, if φ is not convex, Moreau-Rockafellar subdifferentials in definition 2.2.2 is a very restrictive notion that cannot be very useful. $\partial_F \varphi(\bar{x})$ characterizes local properties of φ near $\bar{x}$. For example Fréchet subdifferentiability implies lower semicontinuity [65, Proposition 1.7].

Definition 2.3.1 of the Fréchet subdifferential can be reformulated in the following way [65, Proposition 1.5].

Proposition 2.3.1. $x^* \in \partial_F \varphi(\bar{x})$ *if and only if there exists a function* $g : X \to R$ *such that*

- $g(u) \leq \varphi(u)$ *for any* $u \in X$ *and* $g(\bar{x}) = \varphi(\bar{x})$*;*
- g *is Fréchet differentiable at* $\bar{x}$ *and* $\nabla g(\bar{x}) = x^*$*.*

In addition, one can consider a Fréchet superdifferential

$$\partial_F^+ \varphi(\bar{x}) = \{x^* \in X^* : \limsup_{x \to \bar{x}} \frac{\varphi(x) - \varphi(\bar{x}) - \langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0\}. \tag{2.9}$$

Clearly

$$\partial_F(-\varphi)(\bar{x}) = -\partial_F^+ \varphi(\bar{x}).$$

Fréchet superdifferential is also weak* closed and convex. While set in Definition 2.3.1 consists of linear continuous functions "supporting" φ from below, functions in (2.9) "support" φ from above. Clearly, when φ is not differentiable, the sets $\partial_F \varphi(\bar{x})$ and $\partial_F^+ \varphi(\bar{x})$ may not non-empty and equal. The following is a result that shows Fréchet subdifferential and superdifferential are nonempty and equal when the function φ is Fréchet differentiable[65, Proposition 1.3].

Proposition 2.3.2. *Let φ be Fréchet differentiable ($\nabla\varphi(\bar{x})$ exists) at $\bar{x}$. In this case, one has*

$$\partial_F\varphi(\bar{x}) = \partial_F^+\varphi(\bar{x}) = \{\nabla\varphi(\bar{x})\}.$$

Note that Gâteaux differentiable functions can be Fréchet nonsubdifferentiable [65, Example 1.4]. Both sets $\partial_F\varphi(\bar{x})$ and $\partial_F^+\varphi(\bar{x})$ can be simultaneously empty [65, Example 1.2]. Indeed, if the set $\partial_F\varphi(\bar{x})$ is a singletone, then may not the differentiability of φ at $\bar{x}$ [65, Example 1.3].

In the nonconvex case, the definition of normal cone in subsection 2.2.2 has been modified to Fréchet normal cone.

Definition 2.3.2. *The Fréchet normal cone $N_F(\bar{x};\Omega)$ to a subset Ω of X at $\bar{x}\in\Omega$ is*

$$x^* \in N_F(\bar{x};\Omega) \Leftrightarrow \limsup_{\substack{x \to \bar{x} \\ x\neq\bar{x}}} \frac{\langle x^*, x-\bar{x}\rangle}{\|x-\bar{x}\|} \leq 0. \tag{2.10}$$

$N_F(\bar{x};\Omega)$ is a weak* closed and convex cone closely related to the subdifferential defined above. Fréchet normal cone generalizes the corresponding notion of convex analysis [65, Proposotion 1.19]. In the following, we state simple properties of Fréchet normal cones for the set Ω at $\bar{x}\in\Omega$ [65, Proposition 1.20, 1.21, 1.22, 1.26, 1.27, 1.29].

(i) $N_F(\bar{x};\Omega) = N_F(\bar{x};\mathrm{cl}\Omega)$.

(ii) Let Ω be a cone. Then

$$N_F(\lambda\bar{x};\Omega) = N_F(\bar{x};\Omega) \text{ for any } \lambda > 0.$$

(iii) Let $\Omega = \Omega_1 \cap \Omega_2$. Then

$$N_F(\bar{x};\Omega_2) + N_F(\bar{x};\Omega_1) \subset N_F(\bar{x};\Omega).$$

(iv) If $\Omega^{'} \subset \Omega$, then

$$N_F(\bar{x};\Omega) \subset N_F(\bar{x};\Omega^{'}).$$

(v) Let $\Omega = \Omega_1 + \Omega_2$, $\bar{x} = x_1 + x_2$, $x_i \in \Omega_i$, i=1,2. Then

$$N_F(\bar{x};\Omega) \subset N_F(x_1;\Omega_1) \cap N_F(x_2;\Omega_2).$$

(vi) Let $X = X_1 \times X_2$, $\Omega = \Omega_1 \times \Omega_2$, $\bar{x} = (x_1, x_2)$, $x_i \in \Omega_i \subset X_i$, i=1,2. Then

$$N_F(\bar{x};\Omega) = N_F(x_1;\Omega_1) \times N_F(x_2;\Omega_2).$$

Fréchet subdifferentials and directional derivatives

Fréchet subdifferentials are defined directly without invoking any local approximations of a function. In the following some kind of directional derivative of nonsmooth functions are considered which is closely related to Fréchet subdifferentials. Subderivative [7, 53, 55, 87, 88, 94] and weak subderivative [60, 97] of function φ at $\bar{x} \in \Omega$ in direction $z \in \Omega$ are defined as follows:

$$d\varphi(\bar{x})(z) = \liminf_{\substack{t\downarrow 0\\ y\to z}} \frac{\varphi(\bar{x}+ty)-\varphi(\bar{x})}{t}, \tag{2.11}$$

$$d_w\varphi(\bar{x})(z) = \liminf_{y\overset{w}{\to}z, t\downarrow 0} \frac{\varphi(\bar{x}+ty)-\varphi(\bar{x})}{t}, \tag{2.12}$$

where $y \overset{w}{\to} z$ means that y tends to z in the weak topology of X. $d\varphi(\bar{x})(z)$ and $d_w\varphi(\bar{x})(z)$ are positively homogenous functions from X into $R \cup \{\pm\infty\}$, lower semincontinuous in the norm and the weak topology of X. The inequality

$$d_w\varphi(\bar{x})(z) \le d\varphi(\bar{x})(z),$$

holds for any $z \in X$. If dim$X < \infty$, then both subderivatives coincide. In general, the functions are different and they can differ from the usual directional derivative. There are conditions on function φ under which the subderivative reduces to the usual directional derivative at $\bar{x}$ in direction z [34, 51, 85].

$d_w\varphi(\bar{x})(z)$ is in a sense the lowest possible directional derivative. The following result shows its close relation with Fréchet subdifferential [65, Propositon 1.17].

Proposition 2.3.3. *The following inclusion holds:*

$$\partial_F\varphi(\bar{x}) \subset \{x^* \in X^* : \ d_w\varphi(\bar{x})(z) \geq \langle x^*, z\rangle \ \ \forall z \in X\}. \tag{2.13}$$

If X is reflexive, then in (2.13) the equality holds.

Approach to relate Fréchet normals with distance function

Another approach to defining Fréchet normal cone is based on considering the subdifferential of distance function. Recall that the distance function to Ω is defined by the formula

$$d_\Omega(\bar{x}) = \inf_{x\in\Omega} \|\bar{x} - x\|.$$

Proposition 2.3.4. *[60, 65]* $\partial_F d_\Omega(\bar{x}) = \{x^* \in N_F(\bar{x}; \Omega) : \ \|x^*\| \leq 1\}$.

The following corollary gives an equivalent definition of Fréchet normal cone.

Corollary 2.3.5. $N_F(\bar{x}; \Omega,) = \{\lambda x^* : \ \lambda > 0, \ x^* \in \partial_F d_\Omega(\bar{x})\}$.

Approach to relate Fréchet subdifferential and Fréchet normal cone

The converse is also true: Fréchet subdifferential of an arbitrary function can be equivalently defined through Fréchet normal cone to its epigraph [65, Proposition 1.31, Corollary 1.31.1].

Corollary 2.3.6. $\partial_F\varphi(\bar{x}) = \{x^* \in X^* : \ (x^*, -1) \in N_F((\bar{x}, \varphi(\bar{x})); epi\varphi)\}$.

2.3.2 Dini-Hadamard subdifferentials and normal cones

Another closely related notion is the notion of directional or contingent subdifferential. This notion was introduced by Bouligand for non differentiable functions [15, 16, 17].

Definition 2.3.3. *Given function $\varphi : X \to \bar{R}$ and finite at $\bar{x}$. Dini-Hadamard or Bouligand or contingent subdifferential of $\varphi : X \to \bar{R}$ at $\bar{x}$ is the set of $\partial_D\varphi(\bar{x})$ of $x^* \in X^*$ satisfying the following property: for every $\bar{x} \in X$ and $\epsilon > 0$ there exists some $\delta > 0$ such that*

$$\varphi(\bar{x} + tu) - \varphi(\bar{x}) - \langle x^*, tu\rangle \geq -\epsilon t, \quad \forall (t, u) \in (0, \delta) \times B(\bar{x}, \delta). \tag{2.14}$$

In other words, $x^* \in \partial_D\varphi(\bar{x})$ if and only if

$$\liminf_{\substack{u \to x \\ t \downarrow 0}} \frac{\varphi(\bar{x} + tu) - \varphi(\bar{x}) - \langle x^*, tu\rangle}{t} \geq 0. \tag{2.15}$$

Thus, one has

$$\partial_F\varphi(\bar{x}) \subset \partial_D\varphi(\bar{x}).$$

Let us consider some elementary properties of the Dini-Hadamard subdifferentials [86, Proposition 4.6, 4.8, 4.9, 4.10, 4.11, 4.12]. Most of these properties are common for Fréchet subdifferentials and Dini-Hadamard subdifferentials [65].

Proposition 2.3.7. *If X is finite-dimensional, then $\partial_F\varphi(\bar{x}) = \partial_D\varphi(\bar{x})$.*

Proposition 2.3.8. *The subdifferentials ∂_F and ∂_D are local in the sense that if f and g coincide on some neighborhood of $\bar{x}$, then $\partial_F f(\bar{x}) = \partial_F g(\bar{x})$ and $\partial_D f(\bar{x}) = \partial_D g(\bar{x})$.*

Proposition 2.3.9. *If φ is a convex function, then $\partial_F\varphi(\bar{x})$ and $\partial_D\varphi(\bar{x})$ coincide with Moreau-Rockafellar subdifferential $\partial\varphi(\bar{x})$:*

$$\partial_F\varphi(\bar{x}) = \partial_D\varphi(\bar{x}) = \partial\varphi(\bar{x}) = \{x^* \in X^* : \ \varphi(x) - \varphi(\bar{x}) \geq \langle x^*, x - \bar{x}\rangle, \ \ \forall x \in X\}.$$

Proposition 2.3.10. *For every function φ and finite at $\bar{x}$, $\partial_F\varphi(\bar{x})$ (resp. $\partial_D\varphi(\bar{x})$) is a closed (resp. weak* closed) convex subset of X^*.*

Proposition 2.3.11. *Let $f \leq g$ and $f(\bar{x}) = g(\bar{x})$ is finite. Then*

$$\partial_F f(\bar{x}) \subset \partial_F g(\bar{x}), \;\; \partial_D f(\bar{x}) \subset \partial_D g(\bar{x}).$$

The following results show the necessary conditions by using Fréchet and Hadamard subdifferentials..

Theorem 2.3.12. *If φ attains a local minimum at $\bar{x}$, then one has $0 \in \partial_F \varphi(\bar{x})$ and $0 \in \partial_D \varphi(\bar{x})$.*

Proposition 2.3.13. *If $f + g$ attains a local minimum at $\bar{x}$ and if f is Fréchet differentiable (resp. Hadamard differentiable) at $\bar{x}$, then*

$$-f^{'}(\bar{x}) \in \partial_F g(\bar{x}) \; (resp - \partial f^{'}(\bar{x}) \in \partial_D g(\bar{x})),$$

Corresponding to Dini-Hadamard subdiferential, Dini-Hadamard normal cone is defined.

Definition 2.3.4. *The Dini-Hadamard normal cone $N_D(\bar{x};\Omega)$ to a subset Ω of X at $\bar{x} \in cl\Omega$ is*

$$x^* \in N_D(\bar{x};\Omega) \Leftrightarrow \limsup_{\substack{u \to x, \bar{x}+tu \in \Omega \\ t \downarrow 0}} \frac{\langle x^*, (\bar{x}+tu) - \bar{x}\rangle}{t} \leq 0. \tag{2.16}$$

It is important to relate normal cone to the notion of subdifferential. In the next result, the subdifferential of indicator function is defined as normal cone in Definitions 2.3.4 and 2.3.2 [86, Proposition 4.13].

Proposition 2.3.14. *For every subset Ω of a normed space X and for every $\bar{x} \in cl\Omega$, the Fréchet normal cone and the directional normal cone to Ω at $\bar{x}$ coincide with the corresponding subdifferentials of the indicator function of Ω: we have, respectively*

$$N_D(\bar{x};\Omega) = \partial_D i_\Omega(\bar{x}) \; and \; N_F(\bar{x};\Omega) = \partial_F i_\Omega(\bar{x}). \tag{2.17}$$

Another characterization of directional subdifferential in terms of directional normal cone for the set epigraph φ is considered in the following proposition [86, Corollary 4.15].

Proposition 2.3.15. *The directional subdifferential $\partial_D\varphi(\bar{x})$ of φ at $\bar{x} \in dom\varphi$ and the normal cone $N_D((\bar{x}, \varphi(\bar{x})), epi\varphi)$ are related by the following equivalence:*

$$x^* \in \partial_D\varphi(\bar{x}) \Leftrightarrow (x^*, -1) \in N_D((\bar{x}, \varphi(\bar{x})); epi\varphi). \tag{2.18}$$

Now, a formulation for optimality criteria based on normal cone is presented.

Proposition 2.3.16. *If φ attains a local minimum at $\bar{x}$ on a subset Ω of X and if φ is Fréchet differentiable, (resp. Hadamard differentiable) at $\bar{x}$, then we have, respectively*

$$-\varphi'(\bar{x}) \in N_F(\bar{x}; \Omega), -\varphi'(\bar{x}) \in N_D(\bar{x}; \Omega).$$

2.3.3 Clarke subdifferentials and normal cones

Another attractive construction of generalized normal cones and subdifferentials is Clarke's directional derivative of locally Lipschitzian functions in Banach spaces which was introduced in [22, 26].

Definition 2.3.5. *Clarke's directional derivative of φ at $\bar{x}$ in the direction $x \in X$, denoted by*

$$\varphi^\circ(\bar{x}; x) = \limsup_{\substack{y \to \bar{x} \\ t \downarrow 0}} \frac{\varphi(y + tx) - \varphi(y)}{t}. \tag{2.19}$$

The function of directions $x \mapsto \varphi^\circ(\bar{x}; x)$ is convex. Motivation for defining this directional derivative was to derive necessary optimality conditions for optimal control problems. The theory of Clarke directional derivative has been applied to a wide variety of problems in analysis, mathematical programming , optimal control and calculus of variations and computation [23, 25, 27, 28, 46, 47, 52, 114].

After introducing Clarke's directional derivative, the corresponding subdifferential of φ at $\bar{x}$ is defined as

$$\partial_C \varphi(\bar{x}) := \{x^* \in X^* : \ \varphi^\circ(\bar{x}; x) \geq \langle x^*, x \rangle\}. \tag{2.20}$$

In the following, we proceed to discuss some of the fundamental results of Clarke generalized gradient [1, Proposition 1,3,4, 6].

- $\partial_C \varphi(\bar{x})$ is a nonempty convex subset of X^*. It is closed in the strong topology of X^* and bounded by Lipschits modulus K; Moreover, $\partial_C \varphi(\bar{x})$ is weak* compact.
- $\varphi^\circ(\bar{x}; x)$ is the support function of $\partial_C \varphi(\bar{x})$. This means that for any $x \in X$, we have

$$\varphi^\circ(\bar{x}; x) = \max\{\langle x, x^* \rangle : \ x^* \in \partial_C \varphi(\bar{x}) \rangle\}. \tag{2.21}$$

- If φ attains a local minimum or maximum at $\bar{x}$, then $0 \in \partial_C \varphi(\bar{x})$.
- If φ is convex, then $\partial_C(\varphi)(\bar{x})$ coincides with the subdifferential in the sense of convex analysis (Definition 2.2.2).
- If φ admits a continuous Gâteaux derivative, then $\partial_C \varphi(\bar{x}) = \{\nabla \varphi(\bar{x})\}$. When $X = R^n$, $\partial_C \varphi(\bar{x})$ reduces to singleton set ξ if and only if φ is convex and Fréchet differentiable and $\nabla \varphi(\bar{x}) = \xi$.

Clarke's directional derivative $\varphi^\circ(\bar{x}; x)$ may not reduce to the classical one $\varphi^{'}(\bar{x}; x)$ even for simple real functions like $\varphi(x) = -|x|$ at $\bar{x} = 0$. Thus it leads to Clarke regularity definition.

Definition 2.3.6. *We say φ is Clarke regular at $\bar{x}$ if for every $x \in X$*

$$\varphi^\circ(\bar{x}; x) = \varphi^{'}(\bar{x}; x). \tag{2.22}$$

Convex functions, continuously differentiable functions and certain quasidifferentiable functions are Clarke regular at every point [26, 70].

Third step of Clarke generalized gradient is definition of normal cone. Clarke normal cone for nonempty closed subset Ω of X at $\bar{x}$ is defined as subdifferential of the distance function $(\partial_C d_\Omega)$ in formula (2.20) where $d_\Omega(\bar{x}) = \inf\{\|\bar{x} - x\| : x \in \Omega\}$. Since d_Ω is locally Lipschitz, its generalized directional derivative as defined in definition 2.3.5 exists. Clarke normal cone denoted by $N_C(\bar{x}, \Omega)$ is a closed cone in X^* generated by $d_\Omega(\bar{x})$. In the next theorem all application of Clarke generalized gradient and normal cone is discussed [24].

Consider the problem of minimizing $\varphi(x)$ subject to $x \in \Omega$, $h_j(x) = 0$ $(i \in I)$ and $g_i(x) \leq 0$ $(j \in J)$, where the functions involved are locally Lipschitz and I, J are finite index sets.

Theorem 2.3.17. *If $\bar{x}$ solves the above problem, then there exist scalars $\lambda > 0$, s_j $(j \in J)$, $r_i \geq 0$ $(i \in I)$ not all zero such that $r_i g_i(x) = 0$ and such that*

$$\partial_C \varphi(\bar{x}) + \sum_{j \in J} s_j \partial_C h_j(\bar{x}) + \sum_{i \in I} r_i \partial_C g_i(\bar{x}) \in -N_C(\bar{x}; \Omega).$$

It is well known that Clarke's subdifferential of Lipschitzian functions is not contained in the corresponding subdifferential in convex analysis sense. For example, Ioffe in [54, 56], shows that in finite dimensional case, Clarke's normal cone is the convex closure of the corresponding approximate normal cone and also approximate normal cone is always smaller than Clarke subdifferential. Clarke subdifferential in many cases is too large and thus it is not very useful in necessary optimality conditions. For example, as the trivial one, consider minimizing $-|x|$ over R; clearly $0 \in \partial_C \varphi(\bar{x})$ while $\bar{x}$ is far removed from the minimum that can be directly detected by other necessary conditions for minimization.

2.3.4 Limiting Fréchet subdifferntials and normal cones

The minimality of the limiting Fréchet subdifferential constructions in comparison to general classes of subdifferentials and Clarke subdifferential is established in [62, 63, 68]. The original nonconvex limiting Fréchet normal cone to closed sets in finite dimensional spaces introduced by Kruger and Mordukhovich [81, 84]

$$\bar{N}(\bar{x};\Omega) = \limsup_{x\to\bar{x}}[cone(x - \sqcap(x;\Omega))] \tag{2.23}$$

where

$$\sqcap(x,\Omega) := \{w \in \Omega : \ \|x - w\| = d(x;\Omega)\}.$$

The corresponding limiting Fréchet subdifferential is defined for lower semi-continuous extended-real-valued functions as:

$$\bar{\partial}\varphi(\bar{x}) := \{x^* \in X^* : \ \ (x^*, -1) \in \bar{N}((\bar{x}, \varphi(\bar{x})); \mathrm{epi}\varphi)\}. \tag{2.24}$$

The initial motivation of mentioned subdifferential and normal cone came from the intention to derive necessary optimality conditions for optimal control problems with endpoint geometric constraints by passing to the limit from free endpoint control problems, which are much easier to handle [66, 81, 84].

Then the concept of limiting Fréchet normal cone was extended to infinite-dimensional Banach spaces as sequential limits of ϵ-normals introduced by Kruger and Mordukhovich [62, 64, 67, 68, 81, 82]. In this section, when we write $x \xrightarrow{\Omega} \bar{x}$, this means $x \to \bar{x}$ and $x \in \Omega$.

Definition 2.3.7. (Generalized normals) *Let Ω be a nonempty subset of X.*

(i) Given $x \in \Omega$ and $\epsilon \geq 0$, define the set of ϵ-normals to Ω at x by

$$N_\epsilon(x,\Omega) := \{x^* \in X^* : \ \limsup_{u\xrightarrow{\Omega}x} \frac{\langle x^*, u - x\rangle}{\|u - x\|} \leq \epsilon\}. \tag{2.25}$$

When $\epsilon = 0$, elements of (2.25) are called Fréchet normals and their collection, denoted by $N_F(x;\Omega)$, is the Fréchet normal cone to Ω at x. If $x \notin \Omega$, we put $N_\epsilon(x;\Omega) := \emptyset$ for all $\epsilon \geq 0$.

(ii) Let $\bar{x} \in \Omega$. Then $x^* \in X^*$ is a limiting Fréchet normal to Ω at $\bar{x}$ if there are sequences

$\epsilon_k \downarrow 0$, $x_k \xrightarrow{\Omega} \bar{x}$, and $x_k^* \xrightarrow{w^*} \bar{x}$ such that $x_k^* \in N_{\epsilon_k}(x_k; \Omega)$ for all $k \in N$. The collection of such normals

$$\bar{N}(\bar{x}; \Omega) := \limsup_{\substack{x \xrightarrow{\epsilon \downarrow 0} \bar{x}}} N_\epsilon(x; \Omega) \tag{2.26}$$

is the limiting Fréchet normal cone to Ω at $\bar{x}$. Put $\bar{N}(\bar{x}, \Omega) := \emptyset$ for $\bar{x} \notin \Omega$.

Limsup in definition 2.3.7 denotes the sequential Kuratowski-Painlevé upper limit of the sets $N_\epsilon(x; \Omega)$.

The following theorem describes the limiting Fréchet normal cone to subsets $\Omega \subset R^n$ that are locally closed around $\bar{x}$ [83, Proposition 1.6]. The latter means that there is a neighborhood U of $\bar{x}$ for which $\Omega \cap U$ is closed.

Theorem 2.3.18. (Limiting normals in finite dimensions) *Let $\Omega \subset R^n$ be locally closed around $\bar{x} \in \Omega$. Then the following representations hold:*

$$\bar{N}(\bar{x}; \Omega) := \limsup_{x \to \bar{x}} N_F(x; \Omega),$$

$$\bar{N}(\bar{x}; \Omega) := \limsup_{x \to \bar{x}} [cone(x - \sqcap(x; \Omega))].$$

The first representation of the limiting normal cone in the previous theorem holds in any Asplund space [83, Theorem 2.35].

Observe that both Fréchet normal cone $N_F(\cdot, \Omega)$ and limiting Fréchet normal cone $\bar{N}(\cdot, \Omega)$ are invariant with respect to equivalent norms on X while the ϵ-normal sets $N_\epsilon(\cdot; \Omega)$ depend on the given norm $\|\cdot\|$ if $\epsilon > 0$. Note also that for each $\epsilon \geq 0$ the sets (2.25) are obviously convex and closed in the norm topology of X^*; hence they are weak* closed in X^* when X is reflexive. In contrast to (2.25), the limiting Fréchet normal cone (2.26) may be nonconvex.

Limiting Fréchet subdifferentials are defined through limiting Fréchet normals to epigraphs.

Definition 2.3.8. *Consider a function $\varphi : X \to \bar{R}$ and a point $\bar{x} \in X$ with $|\varphi(\bar{x})| < +\infty$.*

The set

$$\bar{\partial}\varphi(\bar{x}) := \{x^* \in X^* : \ \ (x^*, -1) \in \bar{N}((\bar{x}, \varphi(\bar{x})); epi\varphi)\} \tag{2.27}$$

is the limiting Fréchet subdifferential of φ at $\bar{x}$, and its elements are limiting Fréchet subgradients of φ at this point. We put $\bar{\partial}\varphi(\bar{x}) := \emptyset$ if $|\varphi(\bar{x})| = \infty$.

In [83, Proposition 1.79] and [62, 63], equality of limiting Fréchet normal cone and limiting Fréchet subdifferential of indicator function is shown. The following is an equivalent definition for limiting Fréchet subdifferentials in Asplund spaces (See Theorem 2.3.18).

Definition 2.3.9. *Let $\varphi : X \to \bar{R}$ is lower semicontinous in neighborhood of $\bar{x}$. A limiting Fréchet subdifferntial is defined*

$$\bar{\partial}_F\varphi(\bar{x}) := \{x^* \in X^* : \ \exists\{x_k\} \subset X, \{x_k^*\} \subset X^* \textit{ such that};$$

$$x_k \xrightarrow{\varphi} \bar{x}, \ x_k^* \xrightarrow{w^*} x^*, \ \textit{and}\, x_k^* \in \partial_F(x_k), k = 1, 2, 3, ...\}. \tag{2.28}$$

The notation $x_k \xrightarrow{\varphi} \bar{x}$ and $x_k^* \xrightarrow{w^*} x^*$ mean respectively $x_k \to \bar{x}$ with $\varphi(x_k) \to \varphi(\bar{x})$ and x_k^* converges to x^* in the weak* topology of X^*. In [94], the elements of 2.28 are referred to as general subgradients.

One can rewrite Definition 2.3.9 in the following way:

$$\bar{\partial}_F\varphi(\bar{x}) = \limsup_{x \xrightarrow{\varphi} \bar{x}} \partial_F\varphi(x).$$

2.3.5 Quasidifferentials

Many different approaches such as directional derivatives and Clarke subdifferentials have been studied to approximate broad classes of nonsmooth functions. Construction of some of these approaches are defined in terms of families of linear functions and some requires a special linearization procedure.

The main tool, gradient, in smooth analysis is a point in dual space; the main tool, subdifferential, in convex analysis a convex set and finally the main tool, quasidifferential, in quasidiffenertial calculus is a pair of convex sets called subdifferential and superdifferential.

Denote by P and Q the set of all continuous sublinear and suplinear functions defined on X, respectively. Let $p \in P$, the the set

$$\underline{\partial} p = \{\mu \in X^* : \ \mu(x) \leq p(x) \ \forall x \in X\}, \tag{2.29}$$

is called the sudifferential of function p.

Let $q \in Q$, the set

$$\overline{\partial} q = \{\nu \in X^* : \ \nu(x) \geq q(x) \ \forall x \in X\}, \tag{2.30}$$

is called the superdifferential of function q.

Both subdifferential and superdifferential are nonempty convex w^* compact subsets of $\mathbf{X}^*$.

Let φ be a function defined on an open set $\Omega \subset \mathbf{X}$. We say φ is quasidifferentiable at some point $\bar{x} \in \Omega$, if the directional derivative of φ, $\varphi^{'}(\bar{x}; x)$ exists and it can be represented as $p+q$ where $p \in P$ and $q \in Q$. The function p and q can be linearized by $\underline{\partial} p$ and $\overline{\partial} q$, respectively:

$$p(x) = \max_{\mu \in \underline{\partial} p} \mu(x);$$

$$q(x) = \min_{\nu \in \overline{\partial} q} \nu(x).$$

Now the optimality condition for quasidifferentiabe is presented in the next result .

Theorem 2.3.19. *Let φ be a quasidifferentiable function at point $\bar{x} \in X$ and K be a convex* *cone in X. Then*

$$\min_{u\in K}\varphi^{'}(\bar{x};u)=0 \text{ if and only if } -\overline{\partial}\varphi(\bar{x})\subset\underline{\partial}(\bar{x})-K^{*}.$$

2.3.6 Augmented subdifferentials and normal cones

In [2, 3], the notion of supporting cone was introduced that led to so-called the weak subdifferentials. To eliminate duality gap in nonconvex programming, an augmented lagrangian is used that is constructed by supporting cones [3, 40, 41]. Later in [57], the concept of augmented dual cone is introduced in Banach spaces and a special class of sublinear functions is defined by using the elements of augmented dual cone. Recentely, these concepts are used in [58, 59] to obtain necessary and sufficient conditions of optimality for a wide range of nonconvex and nonsmooth problems in the Euclidean space.

In most cases by using different subdifferentials and normal cones only the necessary part of the optimality condition could be obtained for a nonconvex case. On the other hand, these generalizations do not satisfy the main property of the classical subgradient for nonconvex functions which is linear continuous functions to support the epigraph of functions from below. For example, let function $\varphi : R \to R$ be defined as $\varphi(x) = -|x|$. Then, Clarkes directional derivative is $\varphi^{\circ}(\bar{x};x) = |x|$ for all x and $\partial_C\varphi(\bar{x}) = [-1, 1]$. A similar interpretation is also valid for limiting Fréchet subdifferential, which is defined for this functions as $\partial_M\varphi(\bar{x}) = \{-1, 1\}$. It is clear the elements of these subdifferentials cannot support the epigraph of functions.

In the following, augmented normal cone and weak subdifferential are defined.

Definition 2.3.10. *Let $\Omega \subset \mathbf{X}$. Augmented normal cone is defined by*

$$N_A(\bar{x};\Omega) := \{(x^*,\alpha)\in X^*\times R:\ \langle x^*, x-\bar{x}\rangle+\alpha\|x-\bar{x}\|\leq 0\ \forall x\in\Omega\}, \tag{2.31}$$

Definition 2.3.11. *Let $\varphi : \mathbf{X} \to R$ be a function. Weak subdifferential of function φ is defined as*

$$\partial_w \varphi(\bar{x}) := \{(x^*, \alpha) \in X^* \times R : \ \varphi(x) - \varphi(\bar{x}) \geq \langle x^*, x - \bar{x} \rangle + \alpha \|x - \bar{x}\| \ \forall x \in \Omega\}. \qquad (2.32)$$

Weak subdifferential is a closed and convex set. A generalization of the necessary and sufficient condition is presented by Kasimbeyli and Mammadov [59, Theorem5, Theorem6] for special class of nonconvex functions and nonconvex sets.

Theorem 2.3.20. *Let $\Omega \subset R^n$ and let $\varphi : \Omega \to R$ be a given function. Assume that $\bar{x} \in \Omega$ is a minimizer of φ over Ω, Ω is cone-shaped at $\bar{x}$ and $\Omega \setminus \bar{x} = \emptyset$. Let the directional derivative $\varphi^{'}(\bar{x}; \cdot)$ of φ at $\bar{x}$ be lower semicontinuous on $\mathbf{K} = cone(\Omega - \bar{x})$ and the following two conditions hold:*

- *there exists $\delta > 0$ such that $\varphi(x) - \varphi(\bar{x}) \geq \delta \varphi^{'}(\bar{x}; x - \bar{x})$ for all $x \in \Omega$,*
- $$\beta(\bar{x}) = inf\{\varphi(\bar{x}; h) : \ h \in \mathbf{K} \cap \mathbf{U}\} > 0.$$

Then,

$$(0, 0) \in \partial_w \varphi(\bar{x}) + N_A(\bar{x}; \Omega).$$

Theorem 2.3.21. *If $(0, 0) \in \partial_w \varphi(\bar{x}) + N_A(\bar{x}; \Omega)$, then $\bar{x} \in \Omega$ is a minimum point of function φ on Ω.*

Similar to Proposition 2.2.7, weak subdifferential is described by classical directional derivative ([59, Theorem 1]).

Theorem 2.3.22. *Let $\Omega \subset R^n$ be starshaped with respect to $\bar{x} \in \Omega$ and let $\varphi : \Omega \to R$ be a given function. Suppose that φ has a directional at $\bar{x}$ in every direction $x - \bar{x}$ with arbitrary $x \in \Omega$ and*

$$\varphi(x) - \varphi(\bar{x}) \geq \varphi^{'}(\bar{x}, x - \bar{x}),\ for\ all\, x \in \Omega.$$

In addition, let the directional derivative $\varphi^{'}(\bar{x};\cdot)$ *of* φ *at* $\bar{x}$ *be lower semicontinuous on* $\mathbf{K} = cone(\Omega - \bar{x})$ *and*

$$inf\{\varphi^{'}(\bar{x};h) :\ h \in \mathbf{K} \cap \mathbf{U}\} > -\infty.$$

Then

$$\varphi^{'}(\bar{x};h) = sup\{\langle x^*, h\rangle + \alpha\|h\| :\ (x^*, \alpha) \in \partial_w \varphi(\bar{x}),\ \alpha \leq 0\},\ \forall h \in \mathbf{K}.$$

2.4 Topical functions

Topical functions have applications in cycle time [44] and discrete event systems [4, 42, 43]. Discrete event systems provide a useful abstraction for modelling a wide variety of systems: digital circuits, communication networks, manufacturing plants and etc.

In [80], topical functions are characterized by the fact that the Fenchel-Moreau conjugate function and the conjugate function of type Lau admit very simple explicit descriptions in ordered Banach spaces. Most of these results have been obtained by Rubinov and Singer in finite dimensional case (see [96, 106]).

Idempotency

The word idempotency [29] signifies the study of semirings in which the addition operation is idempotent: $a + a = a$. The best-known example is the max-plus semiring, consisting of the real numbers with negative infinity adjoined in which addition is defined as $\max(a, b)$ and multiplication as $a+b$, the latter being distributive over the former. Interest in such structures arose in the late 1950s through the observation that certain problems of discrete optimisation could be linearised over suitable idempotent semirings. More recently the subject has established intriguing connections with automata theory, discrete event systems, nonexpansive mappings, nonlinear partial differential equations, optimisation theory and large deviations.

In the papers [107, 108] the functions defined on a b-complete idempotent semimodule $\mathbf{X}$ over a b-complete idempotent semifield $\mathbf{K} = (\mathbf{K}, \bigoplus, \bigotimes)$, with values in $\mathbf{K}$, where $\mathbf{K}$ may (or may not) contain a greatest element $\sup \mathbf{K}$, and the residuation $\frac{x}{y}$ is not defined for $x \in \mathbf{X}$ and $y = \inf \mathbf{X}$. In [108, 109], it has been assumed that $\mathbf{K}$ does not have any greatest element, then the greatest element $\tau = \sup \mathbf{K}$ is added to $\mathbf{K}$ and also the operations $\bigoplus$ and $\bigotimes$ in $\mathbf{K}$ are extended to $\bar{\mathbf{K}} := \mathbf{K} \cup \tau$, therefore, for any $x \in \mathbf{X}$, obtain a meaning for $x/\inf \mathbf{X}$, and study functions with values in $\mathbf{K}$. In fact we consider two different extensions of the product $\bigotimes$ from $\mathbf{K}$ to $\bar{\mathbf{K}}$, denoted by $\bigotimes$ and $\dot{\bigotimes}$ respectively, and use them to give characterizations of topical functions.

2.5 Infinite horizon optimization

Many important planning problems such as capacity expansion, equipment replacement and production planning involve sequences of related decisions over an infinite time horizon. The mathematical formulation of such problems leads to infinite horizon optimization which is the optimization problem of selecting an infinite sequence of decisions such that the associated cost over an unbounded horizon is minimum [9, 10, 33, 35, 74, 75, 90, 91, 100].

In many studies, an optimal solution/trajectory of an infinite horizon optimization problem is approximated by a sequence of finite horizon optimal solutions [10, 100, 101, 102, 103]. The approximation of the infinite horizon problem with a finite horizon problem may lead to error; however, it may be possible to establish a bound on the error (See [8, 71]).

The uniqueness of optimal solution is a common assumption used in many studies [9, 10, 49]. In discrete decision problems, unique optimal may be difficult to meet (See [49]). Moreover, discrete decisions arise in many problems including production planning, capacity expansion and equipment replacement. Among the studies that do not use the uniqueness assumption we mention [49, 91, 100, 101].

Along with uniqueness assumption, other conditions are needed to consider the convergence of finite optimal horizons to the optimal solution. For example, in [10], a condition called "weak reachability" is stated that is both necessary and sufficient for a general algorithm to converge to optimal solution.

Total cost

Since the cost over an unbounded horizon may be infinite, a discounting factor is applied in the definition of the total cost. In this case, the infinite horizon problem is to find a sequence of policy decisions that minimizes the discounted costs over the infinite horizon [9, 10]. It is clear that even in the presence of discounting, the total cost may still be infinite. In this case, different optimality criteria apart from minimal total cost are required [19, 69, 98, 99, 102]. Efficiency or finite optimality [45, 91, 102], the average cost [12, 38, 113], overtaking optimality [19, 39, 72, 115, 116] and 1-optimality [13, 112] are some examples of such optimality criteria.

Average cost is one of the commonly used criterion when the total cost is infinite and obtained by replacing the original cost by its average value. One of the the important disadvantages of average optimality is that the average value of an infinite horizon cost is insensitive to the costs incurred over any finite horizon. This behaviour is correctable in the stationary case by restricting consideration to a set of strategies within which average cost minimization makes sense. In order to extend the average cost to nonstationary problems, efficiency concept are used [102]. A feasible solution is said to be efficient or finite optimal if it reaches each of the states through which it passes at minimum cost [45, 91, 100]. It has been shown that efficient solutions always exists and that, under a state reachability condition, the efficient solutions are also optimal average [102]. In [113], a nonhomogeneous stochastic infinite horizon optimization problems whose objective is to minimize the overall average cost is considered. The stochastic problem is transformed into the deterministic problem in order to be able to show the optimal solution exists and is average optimal. The theory is applied to nonhomogeneous infinite horizon Markov decision process. Average cost optimality in the homogenous case has been extensively studied in [31, 36, 89, 95, 111].

In time invariant and periodic control systems which operate on an infinite time horizon, optimal cost functionals are unbounded as time tends to infinity. In order to define new optimality, several attempts are made in this direction; for instance in [39, 115], the notion of overtaking optimality are developed.

In [72], time invariant and periodic control systems which have unbounded optimal cost as time tends to infinity is considred. It is shown under a controllability type condition, that a linear expression can be subtracted from the cost functional, it reduces to bounded costs. Particularly, the existence of overtaking optimal solutions for control systems whose cost contain a discounting factor is established.

In [13, 50], a special case of the general dynamic programming problem, has been solved. The problem is to choose a policy which maximize the total expected income where the total expected income for the policy π is formulated as

$$V(\pi) = \sum_{n=0}^{\infty} \beta^n Q_n(\pi) r(f_{n+1}),$$

where

- β is a discounting factor in $[0,1)$;
- Q_n is a Markov matrix;
- f_{n+1} is a function from the set of states to the set of actions and
- $r(f)$ is a column vector whose sth element is the immediate income $i(s, f(s))$ where s is in the set of states.

For the case $\beta = 1$, the total income from a given policy is typically infinite.

A policy that maximizes the total expected income is called $\beta-$optimal for any $0 \leq \beta < 1$. For the case $\beta = 1$, a policy called 1-optimal if the difference between the total expected discounted return with that policy and the $\beta-$optimal policy for $0 \leq \beta < 1$ tends to 0 when β tends to 1.

Chapter 3

Necessary and sufficient conditions for local optimality via weak subdifferentials

3.1 Relation between weak subdifferentials and Hadamard lower directional derivatives

In this section, necessary and sufficient optimality condition for functions with Hadamard directional derivative are presented. Then, by applying this optimality condition, we show that Hadamard lower directional derivative of a function can be expressed as the supremum of weak subgradients of the function.

Paper:

RELATION BETWEEN WEAK SUBDIFFERENTIALS AND HADAMARD LOWER DIRECTIONAL DERIVATIVES

Published by:

The 44^{th} Annual Iranian Mathematics Conference, August 2013

Authors: Sara Hassani[a, b] and Musa Mammadov[a, b]

[a]Centre for Informatics and Applied Optimization,

School of Science, Information Technology and Engineering,

University of Ballarat, VIC 3353, Australia

[b]National ICT Australia, VRL, VIC 3010, Australia

Corresponding author:

Sara Hassani

e-mail: sarahassani@students.ballarat.edu.au

The Extended Abstracts of Posters of
The 44$^{\text{th}}$ Annual Iranian Mathematics Conference
27-30 August 2013, Ferdowsi University of Mashhad, Iran.

RELATION BETWEEN WEAK SUBDIFFERENTIALS AND HADAMARD LOWER DIRECTIONAL DERIVATIVES

SARA HASSANI[1,2]* AND MUSA MAMMADOV[1,2]

[1] *Center for Informatics and Applied Optimization, School of Science, Information Technology and Engineering, University of Ballarat, VIC 3353, Australia.*
sarahassani@students.ballarat.edu.au
m.mammadov@ballarat.edu.au

[2] *National ICT Australia, VRL, Melbourne, Australia*

ABSTRACT. In this paper, the relation between weak subdifferentials and functions which have Hadamard directional derivative is studied. The optimality condition for Hadamard directional differentiable functions are considered.

1. INTRODUCTION

The notion of subdifferential plays an important role in optimization theory. It was first introduced as a generalization of the concept of ordinary derivative to deal with optimization problems involving convex and nonsmooth functions. Various kinds of subdifferentials have been introduced which are applicable for different classes of nonconvex nonsmooth optimization problems.

Weak subdifferentials are one of the most useful global concepts in terms of deriving the necessary and sufficient conditions of optimality

2010 *Mathematics Subject Classification.* Primary 65K10; Secondary 49K99.
Key words and phrases. weak subdifferential, Hadamard lower directional derivative, optimality condition.

* Speaker.

for a wide range of nonconvex and nonsmooth problems. In this paper $\Omega \subset R^n$. The definition of weak subdifferential is as follows [1]:

Definition 1.1. A pair of $(x^*, \alpha) \in R^n \times R$ is called a weak subgradient of f at $\bar{x}$ on Ω if

$$f(x) - f(\bar{x}) \geq \langle x^*, x - \bar{x} \rangle + \alpha \|x - \bar{x}\| \ \ \forall x \in \Omega. \tag{1.1}$$

The set

$$\partial_\Omega^w f(\bar{x}) = \{(x^*, \alpha) \in R^n \times R : \ (1.1) \text{ is satisfied}\}$$

of all weak subgradients is called the weak subdifferential of f at $\bar{x}$ on Ω.

In [2], the relation of weak subdifferentials and the functions which have directional derivative is studied. Note that the directional derivative of function f at the point $\bar{x}$ in the direction v is as:

$$f^{'}(\bar{x}; v) = \lim_{t \downarrow 0} \frac{f(\bar{x} + tv) - f(\bar{x})}{t}.$$

In this paper, we consider a broader class of functions that have Hadamard directional derivative and we establish the relation with weak subdifferentials. Following is the definition of Hadamard lower directional derivative:

Definition 1.2. Let function $f : R^n \to R$ be finite at $\bar{x} \in R^n$. The Hadamard lower directional derivative at the point $\bar{x} \in R^n$ in the direction $v \in R^n$ is defined by

$$f^H(\bar{x}; v) = \liminf_{\substack{u \to v \\ t \downarrow 0}} \frac{f(\bar{x} + tu) - f(\bar{x})}{t}.$$

2. Main results

In the next theorem, the optimality condition of functions with Hadamard lower directional derivative is given.

Lemma 2.1. *Let $\Omega \subseteq R^n$ with $\bar{x} \in \Omega$ and let $f : \Omega \to R$ be a given function. Suppose that f has a Hadamard lower directional derivative at $\bar{x}$ in every direction $x - \bar{x}$ with arbitrary $x \in \Omega$ and*

$$f(x) - f(\bar{x}) \geq f^H(\bar{x}; x - \bar{x}) \ \ \forall x \in \Omega. \tag{2.1}$$

Then, $\bar{x} \in \Omega$ is a point of minimum of f over Ω, if and only if

$$f^H(\bar{x}; x - \bar{x}) \geq 0, \ \forall x \in \Omega. \tag{2.2}$$

Proof. Let (2.2) be satisfied. Then, by (2.1) we have $f(x) - f(\bar{x}) \geq 0$, $\forall x \in \Omega$ which implies that $\bar{x} \in \Omega$ is a minimal point of f over Ω.

Now let $\bar{x} \in \Omega$ be a minimal point of f over Ω. Then $\forall x \in \Omega$, $f^H(\bar{x}; x - \bar{x})$ exists that means there exist $t_n \downarrow 0$ and $u_n \to x - \bar{x}$ provided that

$$\begin{aligned} f^H(\bar{x}; x - \bar{x}) &= = \liminf_{\substack{u \to x - \bar{x} \\ t \downarrow 0}} \frac{f(\bar{x} + tu) - f(\bar{x})}{t} \\ &= \lim_{\substack{u_n \to x - \bar{x} \\ t_n \downarrow 0}} \frac{f(\bar{x} + t_n u_n) - f(\bar{x})}{\lambda_n}. \end{aligned} \tag{2.3}$$

As $\bar{x}$ is the minimum point, by equation (2.3), we have:

$$f^H(\bar{x}; x - \bar{x}) \geq 0.$$

$\square$

The following theorem describes an important property of the weak subdifferentials when function is Hadamard lower directional differentiable. When a function has directional derivative, we have $f^{'}(\bar{x}, 0) = 0$ but this equality doesn't hold for Hadamard directional derivative. We set $f^H(\bar{x}; 0) = 0$. This condition according to [3, Proposition 4.4], is equivalent to some properties of f such as f is calm at $\bar{x}$. In the following, U is unit sphere and $cone(\cdot)$ stands for the minimal cone containing the enclosed set $\Omega \subset R^n$.

Theorem 2.2. *Suppose that all the conditions of Lemma 2.1 are satisfied. Let the Hadamard lower directional derivative $f^H(\bar{x}; x - \bar{x})$ of f at $\bar{x}$ be lower semicontinuous on $K = cone(\Omega - \bar{x})$ and*

$$inf\{f^H(\bar{x}; h) : h \in K \cap U\} > -\infty. \tag{2.4}$$

In addition, let $f^H(\bar{x}; 0) = 0$. Then f is weakly subdifferentiable at $\bar{x}$ on Ω; that is, $\partial^w_\Omega f(\bar{x}) \neq \emptyset$ and

$$\begin{aligned} &f^H(\bar{x}; h) = \\ &\sup\{\langle x^*, h\rangle + \alpha\|h\| : \ (x^*, \alpha) \in \partial^w_\Omega f(\bar{x}), \alpha \leq 0\} \ \ \forall h \in K. \end{aligned} \tag{2.5}$$

Proof. For the sake of simplicity we denote

$$\phi(h) = f^H(\bar{x}; h), \ \ \forall h \in K.$$

Clearly $\Phi(h)$ is positively homogeneous and $\phi(0) = 0$. By assumption (2.4), $\phi(h)$ is bounded from below on $K \cap U$. Then it is clear that, given any $x^* \in R^n$ the relation

$$\phi(h) \geq \langle x^*, h \rangle + \alpha \|h\| \ \ \forall h \in K \cap U$$

is satisfied for sufficiently small number $\alpha \to -\infty$. Let (x^*, α) satisfies this relation. Since ϕ is positively homogeneous, the above relation is also satisfied for all $h \in K$, and in particular

$$\phi(h) - \phi(0) \geq \langle x^*, h \rangle + \alpha \|h\| \ \ \forall h \in \Omega - \bar{x}. \tag{2.6}$$

This means that $(x^*, \alpha) \in \partial^w_{\Omega \setminus \bar{x}} \phi(0)$; that is, ϕ is weakly subdifferentiable on $\Omega \setminus \bar{x}$ at 0 : $\partial^w_{\Omega \setminus \bar{x}} \phi(0) \neq \emptyset$. Now we show that

$$\partial^w_{\Omega} f(\bar{x}) = \partial^w_{\Omega \setminus \bar{x}} \phi(0). \tag{2.7}$$

Let $(x^*, \alpha) \in \partial^w_{\Omega \setminus \bar{x}} \phi(0)$. Then from (2.1) and (2.6) it follows that (1.1) is satisfied; that is, $(x^*, \alpha) \in \partial^w_{\Omega} f(\bar{x})$.

If $(x^*, \alpha) \in \partial^w_{\Omega} f(\bar{x})$ then for any fixed $x \in \Omega$ we have

$$\phi(x - \bar{x}) = f^H(\bar{x}; x - \bar{x}) = \lim_{\substack{u_n \to x - \bar{x} \\ t_n \downarrow 0}} \frac{f(\bar{x} + t_n u_n) - f(\bar{x})}{t_n} \geq$$

$$\lim_{\substack{u_n \to x - \bar{x} \\ t_n \downarrow 0}} \frac{\langle x^*, t_n u_n \rangle + \alpha \|t_n u_n\|}{t_n} = \langle x^*, x - \bar{x} \rangle + \alpha \|x - \bar{x}\|$$

that leads to (2.6), that is $(x^*, \alpha) \in \partial^w_{\Omega \setminus \bar{x}} \phi(0)$. Then (2.7) is true which means that f is weakly subdifferentiable at $\bar{x}$ on Ω.

The proof of the equality (2.5) is the same as [2, Theorem 1]. $\square$

References

1. A.Y. Azimov and R.N. Gasimov, *On weak conjugacy, weak subdifferentials and duality with zero gap in nonconvex optimization*, International Journal of Applied Mathematics, 1 (1999), no. 4, 171–192.
2. R. Kasimbeyli and M. Mammadov, *Optimality conditions in nonconvex optimization via weak subdifferentials*, Nonlinear Analysis: Theory, Methods and Applications, 74 (2011), no. 7, 2534–2547.
3. J.P. Penot, *Calculus Without Derivatives*, Graduate texts in mathematics, Springer New York, 2013.

3.2 Some geometrical properties of non-convex sets in finite dimensions

This section introduces the notion of $\sigma-$suporting cone based on the definition of augmented normal cone. By using this cone, two new concepts, conic gap and maximal conic gap are defined for a set at a point. These concepts are applied to investigate if a set has a conic gap at some boundary points and moreover to measure how big such a gap is. Finally, the relation of maximal conic gap with $\sigma-$suporting cone and Fréchet normal cone is given.

Paper:

Some geometrical properties of non-convex sets in finite dimensions

Published by:

The 45^{th} Annual Iranian Mathematics Conference, August 2014

Authors: Sara Hassani[a, b], Musa Mammadov[a, b] and Mina Jamshidi[c]

[a]School of Science, Federation University Australia

[b]National ICT Australia, VRL, Melbourne, Australia

[c]Graduate University of advanced technology, Kerman, Iran

Corresponding author:

Sara Hassani

e-mail: Sara.Hassani@nicta.com.au

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین
کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

Some geometrical properties of non-convex sets in finite dimensions

S. Hassani [a] , Ph. D. student at School of Science,
Federation University Australia, Sara.Hassani@nicta.com.au
M. Mammadov[a], Academic member at School of Science
Federation University of Australia, m.mammadov@federation.edu.au
[a]National ICT Australia, VRL, Melbourne, Australia
M. Jamshidi,* Academic member at graduate university of advanced technology,
Graduate University of advanced technology, Kerman, Iran, m.jamshidi@kgut.ac.ir.

Abstract: In this paper, the notions of σ-supporting cone and maximal conic gap are introduced. They are based on the concept of "augmented normal cones". We characterize non-convex sets in finite dimensional space R^n by applying the notion of maximal conic gap. The relationship between the maximal conic gap, $\sigma-$supporting cones and Fréchet normal cones is investigated.
Keywords: σ-supporting cone, maximal conic gap, normal cone.

1 Introduction

The study of geometrical properties of sets is an important problem in many areas of mathematics. For example convexity and non-convexity of sets have significant roles in studying necessary and sufficient conditions of optimality in optimization theory.

Augmented normal cones are useful global concepts in terms of deriving necessary and sufficient conditions of optimality for a wide range of non-convex and non-smooth problems [3, 1].

In this paper, we suppose that Ω is subset of R^n and $\bar{x} \in \Omega$. Tangent cone and Fréchet normal cone are defined as follows:

$$T(\bar{x};\Omega) := \{w \in R^n : \exists\, x_k \xrightarrow{\Omega} \bar{x}, \tau_k \downarrow 0 \ s.t \ \tfrac{x_k - \bar{x}}{\tau_k} \to w\},$$

$$N^F(\bar{x};\Omega) := \{x^* \in R^n : \limsup_{\substack{x \xrightarrow{\Omega} \bar{x} \\ x \neq \bar{x}}} \tfrac{\langle x^*, x-\bar{x}\rangle}{\|x-\bar{x}\|} \leq 0\}.$$

$N^a(\bar{x};\Omega)$ denotes the augmented normal cone to the set Ω at $\bar{x}$ [3, 1]:

$$N^a(\bar{x};\Omega) := \{(x^*, \alpha) \in R^n \times R : \\ \langle x^*, x-\bar{x}\rangle + \alpha\|x-\bar{x}\| \leq 0 \quad \forall x \in \Omega\}.$$

Adopting the structure of augmented normal cones, we aim to generalize them by introducing a new local definition which is called "*σ-supporting cone*".

In the following, S is unit sphere and cone stands for the conic hull. Below we suppose that the norm is Euclidean.

*Corresponding Author

Definition 1.1. *σ-supporting cone is denoted by $N^{\sigma}(\bar{x};\Omega)$ and is defined as follows:*

$$N^{\sigma}(\bar{x};\Omega) := cone\{x^* \in S:\ \sigma_{\Omega}(x^*,\bar{x}) < 1\}$$

where

$$\sigma_{\Omega}(x^*;\bar{x}) := \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle x^*, x-\bar{x}\rangle}{\|x-\bar{x}\|}. \tag{1}$$

As an application of σ-supporting cone, a new concept "maximal conic gap" is introduced in order to characterize non-convex sets. The main question of interest here is to investigate if a set has a "conic gap" at some boundary points. Moreover, to measure how "big" such a gap is, we introduce "maximal conic gap".

Definition 1.2. *We say Ω has a conic gap at $\bar{x}$ if there exists $x^* \in S$ such that*

$$\sigma_{\Omega}(x^*,\bar{x}) < 1.$$

Definition 1.3. *The maximal conic gap β^* with respect to Ω at $\bar{x}$ is defined as follows:*

$$\beta^* := -\inf_{x^*\in S}\sigma_{\Omega}(x^*,\bar{x}) = -\inf_{x^*\in S}(\limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle x^*, x-\bar{x}\rangle}{\|x-\bar{x}\|}). \tag{2}$$

The minus sign in (2), indicates that the value (β^*) increases with the size of the gap. It means if $\beta_1^* > \beta_2^*$, then the gap with the value β_1^* is bigger than the other one. Since norm is Euclidean, clearly $-1 \le \beta^* \le 1$.

2 Main results

First we give an example of two sets with different maximal conic gap values.

Example 2.1. *Let $\Omega_1 = \{(x,y) \in R^2 :\ y \ge 0\}$, $\Omega_2 = \{(x,y) \in R^2 :\ y \ge x$ or $y \ge -x\}$ and $\bar{x} = (0,0)$. By Definition 1.3, $\beta_1^* = 0$ and $\beta_2^* = -\frac{1}{\sqrt{2}}$. As the maximal gap values show, the size of the gap in Ω_1 is bigger than Ω_2.*

In the following lemma, we give another representation for $\sigma_{\Omega}(x^*,\bar{x})$ defined in (1).

Lemma 2.2. $\sigma_{\Omega}(x^*,\bar{x}) = \sup\{\langle x^*, z\rangle :\ z \in T(\bar{x};\Omega)\cap S\}$.

Proof. Clearly by definition of $T(\bar{x};\Omega)$, we have:

$$\sup\{\langle x^*, z\rangle :\ z \in T(\bar{x};\Omega)\cap S\} =$$

$$\sup\{\langle x^*, \lim_{k\to\infty}\frac{x_k-\bar{x}}{\tau_k}\rangle :\ x_k \xrightarrow{\Omega} \bar{x}, \tau_k \downarrow 0, \lim_{k\to\infty}\frac{x_k-\bar{x}}{\tau_k} = z,$$

$$\|z\| = 1\} = \sup\{\langle x^*, \lim_{x_k \xrightarrow{\Omega} \bar{x}}\frac{x_k-\bar{x}}{\|x_k-\bar{x}\|}\rangle :$$

$$\lim_{x_k \xrightarrow{\Omega} \bar{x}}\frac{x_k-\bar{x}}{\|x_k-\bar{x}\|} \text{ exists}\}$$

which means $\sigma_{\Omega}(x^*,\bar{x}) = \sup\{\langle x^*, z\rangle :\ z \in T(\bar{x};\Omega)\cap S\}$. ☐

In the following theorem, we give a characterization of $T(\bar{x};\Omega)$ by maximal conic gap value.

Theorem 2.3. (1) *If $-1 < \beta^* < 0$, then $T(\bar{x};\Omega)$ is non-convex.*

(2) *$\beta^* = -1$ if and only if $T(\bar{x};\Omega) = X$.*

Proof. (1). By contradiction let $T(\bar{x};\Omega)$ be convex. By the separation theorem 4.1 in [2], there is a half-space $H^{0-}_{x^*}$ such that

$$T(\bar{x};\Omega) \subseteq H^{0-}_{x^*} = \{x \in R^n : \langle x^*, x\rangle \le 0\}.$$

So

$$\sigma_{\Omega}(x^*,\bar{x}) = \sup\{\langle x^*, z\rangle : z \in T(\bar{x};\Omega)\cap S\} \le 0$$

which results in $\beta^* \ge 0$. This contradicts the relation $-1 < \beta^* < 0$.

(2). Let $\beta^* = -1$. We show that $T(\bar{x};\Omega) = X$. By the definition of β^*, for any $x^* \in S$,

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین
کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

$$\sup\{\langle x^*, z\rangle : z \in T(\bar{x}; \Omega) \cap S\} = 1.$$

Since $T(\bar{x}; \Omega) \cap S$ is closed and bounded and the norm is Euclidean, $x^* \in T(\bar{x}; \Omega)$, and finally since $T(\bar{x}; \Omega)$ is a cone, $T(\bar{x}; \Omega) = X$. □

In the following, the relationship of the maximal conic gap with the σ-supporting cone and Fréchet normal cone is given.

Lemma 2.4. $N^F(\bar{x}; \Omega) \subset N^\sigma(\bar{x}; \Omega)$.

Proof. It is clear by definition $\sigma-$supporting cone. □

Theorem 2.5. (1) *If* $\beta^* > 0$ *then* $intN^F(\bar{x}; \Omega) \neq \emptyset$ *and* $intN^\sigma(\bar{x}; \Omega) \neq \emptyset$.

(2) *If* $\beta^* < 0$ *then* $N^F(\bar{x}; \Omega) = \{0\}$ *and if* $\beta^* > -1$ *then* $N^\sigma(\bar{x}; \Omega) \neq \emptyset$.

Proof. (1) Let $\beta^* > 0$. It means that for some $x^* \in S$, $\sigma_\Omega(x^*, \bar{x}) < 0$. Let $\sigma_\Omega(x^*, \bar{x}) = \alpha$ and $\epsilon = -\frac{\alpha}{2}$. We show that $B_\epsilon(x^*) \subset N^F(\bar{x}; \Omega)$ where $B_\epsilon(x^*)$ is a neighborhood of x^* with ϵ radius.

Let $y^* \in B_\epsilon(x^*)$. We have $y^* = x^* + y^{'}$ for some $y^{'}$ with $\|y^{'}\| \leq \epsilon$ and:

$$\langle x^* + y^{'}, z\rangle \leq \alpha + \|y^{'}\| \leq \alpha - \tfrac{\alpha}{2} < 0 \text{ for all } z \in T(\bar{x}; \Omega) \cap S,$$

therefore $y^* \in N^F(\bar{x}; \Omega)$, and the interior of the Fréchet normal cone is not empty. By lemma 2.4, $intN^\sigma(\bar{x}; \Omega) \neq \emptyset$.

(2) Let $\beta^* < 0$. It means that, for any $x^* \in S$,

$$\limsup_{x \xrightarrow{\Omega} \bar{x}} x \xrightarrow{\Omega} \bar{x} \frac{\langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} > 0.$$

Therefore $N^F(\bar{x}; \Omega) = \{0\}$.

Now let $-1 < \beta^*$. By definition of β^*, we have:

$$-1 < -\inf_{x^* \in S} \sigma_\Omega(x^*, \bar{x}),$$

which means that there exists $x^* \in S$ such that

$$\sigma_\Omega(x^*, \bar{x}) < 1,$$

therefore, $x^* \in N^\sigma(\bar{x}; \Omega)$. □

References

[1] A.Y. Azimov and R.N. Gasimov, *On weak conjugacy, weak subdifferentials and duality with zero gap in nonconvex optimization*, International Journal of Applied Mathematics, 1 (1999), no. 4, 171–192.

[2] L.D. Berkovitz, *Convexity and Optimization in* R^n, Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts, Wiley, 2003.

[3] R. Kasimbeyli and M. Mammadov, *Optimality conditions in nonconvex optimization via weak subdifferentials*, Nonlinear Analysis: Theory, Methods and Applications, 74 (2011), no. 7, 2534–2547.

3.3 Necessary and sufficient conditions for local optimality via weak subdifferentials

The main focus of this section is on optimality conditions for local optimality in finite dimensional normed spaces. We introduce a new local supporting function and then consider the structure of contingent cone of a set around a boundary point of the set by applying this function. The optimality condition for a special class of optimization problems is presented by using the local supporting functions together with weak subdifferentials.

Paper:

Necessary and sufficient conditions for
Local optimality via weak subdifferentials

Published by:

Advances in Theoretical and Applied Mathematics, 2014

Authors: S.Hassani[a, b] and M.A.Mammadov[a, b]

[a]Federation University Australia, Victoria 3353, Australia

[b]National Information and Communications Technology Australia (NICTA)

Corresponding author:

Sara Hassani

e-mail: sarahassani@students.federation.edu.au

Advances in Theoretical and Applied Mathematics.
ISSN 0973-4554 Volume 9, Number 2 (2014), pp. 143–154

http://www.ripublication.com/atam.htm

Necessary and sufficient conditions for Local optimality via weak subdifferentials

S. Hassani
Federation University Australia,
Victoria 3353, Australia
National Information and
Communications Technology Australia (NICTA)
E-mail: sarahassani@students.federation.edu.au

M.A. Mammadov
Federation University Australia,
Victoria 3353, Australia
National Information and
Communications Technology Australia (NICTA)
E-mail: m.mammadov@federation.edu.au

Abstract

In this paper necessary and sufficient conditions for local optimality in finite dimensional normed spaces in terms of weak subdifferential are studied. It is based on newly introduced local sigma supporting function that describes the structure of given set at a particular point by using the contingent cone. In this way, the local optimality versions of related results in [8] are established.

AMS Subject Classification: 65K10.
Keywords: Non-convex analysis, optimality conditions, supporting cone, non-convex sets.

1. Introduction

In this paper the problem of necessary and sufficient conditions of optimality in non-convex optimization problems in a finite dimensional normed space is considered. As a classic technique, Fréchet differentibility helps to derive necessary conditions for optimization problems, however, it can not be applied for deriving sufficient conditions.

Investigations of normal cones to convex sets can be traced back to Minkowski [11]. Later, the concept of global normal cone was treated by Fenchel [4] as the outward normals to supporting half spaces to the set. Deriving necessary and sufficient conditions for Fréchet differentiable functions can be obtained by using Fréchet differentibility and normal cone defined for convex sets.

As the functions are not always Fréchet differentiable, the main approach to derive necessary conditions of optimality for nonsmooth and non-convex problems, is to generalize the notion of differentiability and normal cones. For the first time Moreau and Rockafellar [15] come up with the concept of subdifferentials as a generalization of the concept of ordinary derivative to deal with optimization problems involving convex nonsmooth (nondifferentiable) functions.

For different classes of Nonsmooth and non-convex optimization problems, the variety of different subdifferentials and normal cones have been introduced. Some of the major concepts related to this paper are mentioned here in brief. These concepts (normal cones, subdifferentials) depend on the characterizations of the objective function as well as properties of the variable space. Fréchet subdifferentials were first introduced for finite dimensions in [2] (under the name "lower semidifferentials").

Clarke [3] introduced one of the important generalization of normal cones beyond convexity of functions and sets based on the generalization of the ordinary directional derivative. The locally Lipschitz functions are investigated by Clarke's directional derivative and its related subdifferential. We mention some of the attractive generalizations of directional derivative: generalized directional derivative (upper subderivative) introduced by Rockafellar [16, 17], lower semiderivative introduced by Penot [13], lower Dini (or Dini- Hadamard) directional derivative introduced by Ioffe [5], and subderivative introduced by Rockafellar and Wets [14]. Mordukhovich and Kruger [12] come up with the concept of the non-convex limiting Fréchet normal cone in finite dimensional spaces, and then the concept is extended to infinite dimensional spaces ([10]).

Augmented normal cones and weak subdifferentials are one of the most useful nonlinear global concepts introduced by Azimov and Gasimov [1]. Recently, Kasimbeyli and Mammadov [7, 8] considered "necessary" and "sufficient" conditions of optimality for a wide range of non-convex and nonsmooth problems in Euclidean space. This is the first generalization obtained in the form of a necessary and sufficient condition for global non-convex optimization problems.

In this paper, we consider local optimality conditions for non-convex and nonsmooth optimization problems by applying augmented normal cones and weak subdifferentials similar to global optimality conditions introduced in [8]. Together with this, we also establish the analogies of these results for a broader class of finite dimensional normed spaces.

2. Local supporting function

Throughout the paper we assume that $\mathbf{X}$ is a finite dimensional space with norm $\|\cdot\|$, $\Omega \subset \mathbf{X}$, $\bar{x} \in \Omega$ and $\mathbf{K} = \mathrm{cl}\,(\mathrm{cone}(\Omega - \bar{x}))$ where "cl" stands for the closure of a set.

"cone($\mathbf{A}$)" for a given set $\mathbf{A}$ stands for

$$\text{cone}(\mathbf{A}) = \{\lambda x :\ \lambda \geq 0,\ x \in \mathbf{A}\}.$$

The classical or Bouligand tangent cone, also called the contingent cone, for the set Ω at the point $\bar{x}$ is denoted by $\mathbf{T}_\Omega(\bar{x})$ and defined as follows:

$$\mathbf{T}_\Omega(\bar{x}) = \{w \in \mathbf{X} :\ \exists\, x^v \to \bar{x} \text{ where } x^v \in \Omega \text{ and } \tau^v \downarrow 0 \text{ such that } \frac{x^v - \bar{x}}{\tau^v} \to w\}.$$

The unit sphere and the unit ball of $\mathbf{X}$ are denoted by $\mathbf{U}$ and $\mathbf{B}$, respectively:

$$\mathbf{U} = \{x \in \mathbf{X} :\ \|x\| = 1\},\quad \mathbf{B} = \{x \in \mathbf{X} :\ \|x\| \leq 1\}.$$

The norm of dual space $\mathbf{X}^*$ is denoted by $\|\cdot\|_*$ where

$$\|\cdot\|_* := \max\{\langle\cdot, x\rangle :\ x \in \mathbf{U}\}$$

and $\langle\cdot,\cdot\rangle$ is the scalar product. Note that we use maximum in definition of $\|\cdot\|_*$ as the space $\mathbf{X}$ is finite dimensional space. The unit sphere and unit ball of dual space of $\mathbf{X}$ are denoted by $\mathbf{U}^*$ and $\mathbf{B}^*$, respectively. The notation $B(x;\varepsilon)$ stands for the ball at center x with radius $\varepsilon > 0$; that is,

$$B(x;\varepsilon) := \{y \in \mathbf{X} :\ \|y - x\| \leq \varepsilon\}.$$

We will also denote by "int", "bd" and "co" the interior, the boundary and the convex hull of a set.

Definition 2.1. Given point $\bar{x} \in \Omega$, function $\sigma_L(\cdot;\bar{x}) :\ U^* \to R$ defined by

$$\sigma_L(x^*;\bar{x}) := \max_{y \in \mathbf{T}_\Omega(\bar{x}) \cap \mathbf{U}} \langle x^*, y\rangle. \tag{2.1}$$

will be called a local supporting function.

Clearly the maximum in (2.1) is attained. Moreover, it is not difficult to verify that $\sigma_L(\cdot;\bar{x})$ is a continuous and convex function defined on $\mathbf{X}^*$. The following lemma is a characterization for non-convex sets by the mean of function $\sigma_L(x^*;\bar{x})$.

Lemma 2.2. Assume that the following relation holds

$$0 < \min_{x^* \in \mathbf{U}^*} \sigma_L(x^*;\bar{x}) < 1. \tag{2.2}$$

Then $\mathbf{T}_\Omega(\bar{x})$ is non-convex and given any $y \in \mathbf{X} \setminus \mathbf{T}_\Omega(\bar{x})$ there exists $\varepsilon, \delta > 0$ such that

$$\text{int}(\text{co}\,[B(\delta y;\varepsilon) \cup \{0\}]) \subset (\mathbf{X} \setminus \mathbf{T}_\Omega(\bar{x})) \setminus (\Omega - \{\bar{x}\}). \tag{2.3}$$

Proof. Let $0 < \min_{x^* \in \mathbf{U}^*} \sigma_L(x^*; \bar{x}) < 1$ and on the contrary assume that $\mathbf{T}_\Omega(\bar{x})$ is convex. If $\mathbf{T}_\Omega(\bar{x}) \neq \mathbf{X}$ then there exists a closed hyperplane such that

$$\mathbf{T}_\Omega(\bar{x}) \subseteq \{x \in X : \langle x^*, x\rangle \leq 0\}$$

for some linear function $x^* \in \mathbf{X}^*$. Then

$$\sigma_L(x^*; \bar{x}) = \max\{\langle x^*, y\rangle : y \in \mathbf{T}_\Omega(\bar{x}) \cap \mathbf{U}\} \leq 0,$$

which is a contradiction. On the other hand, if $\mathbf{T}_\Omega(\bar{x}) = \mathbf{X}$ then $\sigma_L(x^*; \bar{x}) = 1$ for all $x^* \in \mathbf{U}^*$ which is again a contradiction. Therefore, $\mathbf{T}_\Omega(\bar{x})$ is non-convex.

Now we show the second assertion of the lemma. Take any

$$y \in \mathbf{X} \setminus \mathbf{T}_\Omega(\bar{x}). \tag{2.4}$$

First we note that there is a sufficiently small number $\delta > 0$ such that

$$\lambda y \notin \Omega - \{\bar{x}\}, \quad \forall \lambda \in (0, \delta]. \tag{2.5}$$

Indeed if $\lambda_k y \in \Omega - \{\bar{x}\}$ for some sequence $\lambda_k \to 0$, then for the sequence $x_k := \lambda_k y + \bar{x}$ we have $[x_k - \bar{x}]/\lambda_k \to y$ or $y \in \mathbf{T}_\Omega(\bar{x})$ that contradicts (2.4).

Denote $z = \delta y$. Since $z \in \mathbf{X} \setminus \mathbf{T}_\Omega(\bar{x})$ and $\mathbf{X} \setminus \mathbf{T}_\Omega(\bar{x})$ is an open cone, there exists a small number $\varepsilon > 0$ such that

$$\mathrm{int}(\mathrm{co}\,[B(z; \varepsilon) \cup \{0\}]) \subset \mathbf{X} \setminus \mathbf{T}_\Omega(\bar{x}).$$

Next we show that the number $\varepsilon > 0$ can be chosen so small that the relation

$$\mathrm{int}(\mathrm{co}\,[B(z; \varepsilon) \cup \{0\}]) \cap (\Omega - \{\bar{x}\}) = \emptyset \tag{2.6}$$

is also satisfied. This will lead to (2.3) and complete the proof of the lemma.

On the contrary assume that (2.6) is not true. Then there are a sequence $\varepsilon_k \to 0$ and a sequence of points $y_k \in \Omega - \{\bar{x}\}$ such that

$$y_k \in \mathrm{int}(\mathrm{co}\,[B(z; \varepsilon_k) \cup \{0\}]), \quad \forall k.$$

Since y_k is bounded, for the sake of simplicity we can assume that $y_k \to y^*$; and clearly $y^* \in \Omega$.

The above relation implies that y_k can be represented in the form

$$y_k = \lambda_k z_k + (1 - \lambda_k)\, 0 = \lambda_k z_k$$

where $\lambda_k \in (0, 1]$ and $z_k \in B(z; \varepsilon_k)$. Clearly, $z_k \to z$ and reminding that $y_k \to y^*$, the sequence λ_k also converges to some number $\lambda^* \in [0, 1]$; that is, $\lambda_k \to \lambda^*$.

Now if $\lambda^* > 0$ then we have $\lambda^* z = \lambda^* \delta y \in \Omega - \{\bar{x}\}$ that contradicts (2.5). Thus, $\lambda^* = 0$ and consequently $\lambda_k \to 0$ and $y_k \to 0$. Denoting $x_k = y_k + \bar{x}$ we obtain $\lim[x_k - \bar{x}]/\lambda_k = \lim z_k = z$ which means that $z \in T_\Omega(\bar{x})$. This again contradicts (2.4).

Lemma is proved. ■

The following example shows that the inverse of Lemma 2.2 is not true; that is, the non-convexity of $T_\Omega(\bar{x})$ does not necessarily means that $\min\limits_{x^* \in \mathbf{U}^*} \sigma_L(x^*; \bar{x}) = 1$.

Example 2.3. Let the set R^2 is equipped with L_∞ (i.e. for any $x \in R^2$, $\|x\|_\infty = \max\{|x_1|, |x_2|\}$). It is clear that the dual norm of R^2 is L_1 (i.e. for any $x \in R^2$, $\|x\|_1 = |x_1| + |x_2|$). Let $\Omega = \{(x, y) \in R^2 : \ y \geq -x \text{ or } y \leq x\}$ and $\bar{x} = (0, 0)$. In this example, $\Omega = T_\Omega(\bar{x})$. It is not difficult to observe that there is no $x^* \in \mathbf{U}^*$ with $\sigma_L(x^*; \bar{x}) < 1$ while $T_\Omega(\bar{x})$ is non-convex.

Below we show that the inverse of Lemma 2.2 is true in strictly convex spaces. Before we prove the related lemma, we present the definition of strictly convex spaces and some properties.

Definition 2.4. (page 112, [18]) Normed space $\mathbf{X}$ is called strictly convex if its unit ball is a strictly convex set; i.e., if $x \neq y$, $x, y \in \mathbf{U}$ and $h = \dfrac{1}{2}(x + y)$ then $\|h\| < 1$.

Let $x' \in \mathbf{U}$. By Theorem 5.20 in [18], there exists $x^* \in \mathbf{U}^*$ such that

$$\langle x^*, x' \rangle = \max_{x \in \mathbf{U}} \langle x^*, x \rangle = 1. \tag{2.7}$$

We also need the following property of strictly convex spaces.

Proposition 2.5. [6] Let $\mathbf{X}$ be strictly convex space and $x^* \in \mathbf{X}^*$. Then the maximum of x^* on unit sphere $\mathbf{U}$ is unique.

The following is a characterization of non-convex sets by applying local supporting function $\sigma_L(x^*; \bar{x})$ in strictly convex spaces which is not true in any normed space as shown in Example 2.3.

Lemma 2.6. Let $\mathbf{X}$ be strictly convex space and $\mathbf{T}_\Omega(\bar{x}) \neq \mathbf{X}$. Then the relation (2.2) holds if and only if $\mathbf{T}_\Omega(\bar{x})$ doesn't belong to any half space; that is, there is no $z^* \in \mathbf{X}^*$ such that $\langle x^*, x \rangle \leq 0$ for all $x \in T_\Omega(\bar{x})$.

Proof. In the proof of Lemma (2.2) it is shown that if (2.2) satisfy then $\mathbf{T}_\Omega(\bar{x})$ doesn't contain in a half space. That is why we consider only the inverse proving that (2.2) holds.

Assume that $\min\limits_{x^* \in \mathbf{U}^*} \sigma_L(x^*; \bar{x}) = 1$. Then $\sigma_L(x^*; \bar{x}) = 1$ for all $x^* \in \mathbf{U}^*$. Take any $x \in \mathbf{U}$. By equation 2.7, there exists $z^* \in \mathbf{X}$ such that

$$\sigma_L(z^*; \bar{x}) = \max_{y \in \mathbf{T}_\Omega(\bar{x}) \cap \mathbf{U}} \langle z^*, y \rangle = 1 = \langle z^*, x \rangle = \max_{y \in \mathbf{U}} \langle z^*, y \rangle.$$

By proposition 2.5, maximum z^* on $\mathbf{U}$ is unique and consequently $x \in \mathbf{T}_\Omega(\bar{x}) \cap \mathbf{U}$. Therefore $\mathbf{T}_\Omega(\bar{x}) = \mathbf{X}$ which is a contradiction. Thus we conclude that there is $x^* \in \mathbf{U}^*$ such that $\sigma_L(x^*; \bar{x}) < 1$.

Now assume that there exists $z^* \in \mathbf{U}^*$ such that

$$\sigma_L(z^*; \bar{x}) = \max_{y \in \mathbf{T}_\Omega(\bar{x}) \cap \mathbf{U}} \langle z^*, y \rangle \leq 0.$$

Then, $\langle z^*, y \rangle \leq 0$ for any $y \in \mathbf{T}_\Omega(\bar{x}) \cap \mathbf{U}$ which means $\mathbf{T}_\Omega(\bar{x})$ contains in a half space. This contradicts the assumption of the lemma.

Lemma is proved. ■

In the last part of this section, we consider a relation between the separation property introduced in [9] and local supporting function $\sigma_L(z^*; \bar{x})$. We start with the definition of separation property.

Definition 2.7. ([9]) Let $\mathbf{C}$ and $\mathbf{K}$ be closed cones of a normed space $\mathbf{X}$. Let $\tilde{\mathbf{C}}$ and $\tilde{K}^\partial$ be the closure of the sets $\mathrm{co}(\mathbf{C} \cap \mathrm{U})$ and $\mathrm{co}((\mathrm{bd}(\mathbf{K}) \cap \mathbf{U}) \cup \{0_\mathbf{X}\})$. The cones $\mathbf{C}$ and $\mathbf{K}$ are said to have the separation property with respect to the norm $\|\cdot\|$ if

$$\tilde{\mathbf{C}} \cap \tilde{K}^\partial = \emptyset. \tag{2.8}$$

Take any positive number $\beta < 1$ and $x^* \in U^*$. Let

$$\mathbf{C} = \mathrm{cone}\{x \in \mathbf{U} : \ \langle x^*, x \rangle \geq \beta\}. \tag{2.9}$$

In the following theorem we show that under some conditions on the local supporting function, the cones $\mathbf{C}$ and $T_\Omega(\bar{x})$ satisfy the separation property.

Theorem 2.8. Let there exists $x^* \in \mathbf{U}^*$ such that $\sigma_L(x^*; \bar{x}) < 1$. Then given any positive number $\beta \in (\sigma_L(x^*; \bar{x}), 1)$, cones $\mathbf{C}$ and $\mathbf{T}_\Omega(\bar{x})$ satisfy the separation property.

Proof. By the assumption of the theorem

$$\max_{y \in \mathbf{T}_\Omega(\bar{x}) \cap \mathbf{U}} \langle x^*, y \rangle = \sigma_L(x^*; \bar{x}) < 1 = \|x^*\|_* = \max_{x \in \mathbf{U}} \langle x^*, x \rangle. \tag{2.10}$$

Denote $\alpha = \sigma_L(x^*; \bar{x})$ and take any $\beta > 0$ such that

$$\alpha = \max_{y \in \mathbf{T}_\Omega(\bar{x}) \cap \mathbf{U}} \langle x^*, y \rangle = \sigma_L(x^*; \bar{x}) < \beta < 1. \tag{2.11}$$

Since $\mathbf{U}$ is closed, there exists $a \in \mathbf{U}$ such that $\langle x^*, a \rangle = \|x^*\|_* = 1$. Then $a \in \mathbf{C}$ and $\mathbf{C} \neq \emptyset$.

Denote $\tilde{\mathbf{C}} = \mathrm{cl}(\mathrm{co}(\mathbf{C} \cap \mathbf{U}))$ and $\tilde{T}_\Omega(\bar{x})^\partial = \mathrm{cl}(\mathrm{co}((\mathrm{bd}(\mathbf{T}_\Omega(\bar{x})) \cap \mathbf{U}) \cup \{0_\mathbf{X}\}))$. We need to prove $\tilde{\mathbf{C}} \cap \tilde{T}_\Omega(\bar{x})^\partial = \emptyset$.

First we show that for any $x \in \tilde{\mathbf{C}}$ the inequality $\langle x^*, x \rangle \geq \beta$ holds. Let $x \in \mathrm{co}(\mathbf{C} \cap \mathbf{U})$. Then the following representation is true $x = \sum_{i=1}^{n+1} \alpha_i x_i$; where $x_i \in \mathbf{C} \cap \mathbf{U}$ and $\sum_{i=1}^{n+1} \alpha_i = 1$. As $x_i \in \mathbf{C} \cap \mathbf{U}$, from (2.9) we have

$$\langle x^*, x \rangle = \sum_{i=1}^{n+1} \alpha_i \langle x^*, x_i \rangle \geq \beta. \tag{2.12}$$

From continuity of $\langle x^*, \cdot\rangle$ and (2.12), for any $x \in \mathrm{cl}(\mathrm{co}(\mathbf{C}\cap\mathbf{U}))$, we have $\langle x^*, x\rangle \geq \beta$.

It is clear from (2.11) that for any $y \in \mathbf{T}_\Omega(\bar{x}) \cap \mathbf{U}$, the relation $\langle x^*, y\rangle \leq \alpha < \beta$ holds. Since $\beta > 0$, we have $\langle x^*, 0\rangle = 0 < \beta$. Thus $\langle x^*, y\rangle \leq \max\{\alpha, 0\} < \beta$ for any $y \in \tilde{T}_\Omega(\bar{x})^\partial$. Therefore $\tilde{\mathbf{C}} \cap \tilde{T}_\Omega(\bar{x})^\partial = \emptyset$. ■

3. Necessary and sufficient conditions of optimality

In ([8]), necessary and sufficient condition of global optimality for a class of non-convex and nonsmooth optimization problems are considered by applying weak subdifferential and augmented normal cone. Below we give the definition of weak subdifferential and augmented normal cone introduced in [8]. Let $f : \Omega \to R$ be a single-valued function. The weak subdifferential of f at $\bar{x}$ on Ω is defined as

$$\partial^w_\Omega f(\bar{x}) = \{(x^*, \alpha) \in X^* \times R : f(x) - f(\bar{x}) \geq \langle x^*, x - \bar{x}\rangle + \alpha\|x - \bar{x}\|, \quad \forall x \in \Omega\}.$$

The set

$$N^A(\bar{x}; \Omega) = \{(x^*, \alpha) \in X^* \times R \ : \ \langle x^*, x - \bar{x}\rangle + \alpha\|x - \bar{x}\| \leq 0, \quad \forall\, x \in \Omega\},$$

is called an augmented normal cone to Ω at $\bar{x}$. These are global concepts and consequently the global optimality condition considered in [8] is

$$(0, 0) \in \partial^w_\Omega f(\bar{x}) + N^A(\bar{x}; \Omega). \tag{3.1}$$

Below we consider the local versions of these definitions, where the set Ω is replaced by

$$\Omega_T(\bar{x}) := (T_\Omega(\bar{x}) + \bar{x}) \cap \Omega.$$

Accordingly, we call corresponding sets $\partial^{lw}_\Omega f(\bar{x})$ and $N^{lA}(\bar{x}; \Omega)$ a local weak subdifferential and a local augmented normal cone, respectively:

$$\partial^{lw}_\Omega f(\bar{x}) = \{(x^*, \alpha) \in X^* \times R : \\ \exists\varepsilon > 0,\ f(x) - f(\bar{x}) \geq \langle x^*, x - \bar{x}\rangle + \alpha\|x - \bar{x}\|,\ \forall x \in \Omega_T(\bar{x}) \cap B(\bar{x}, \varepsilon)\};$$

$$N^{lA}(\bar{x}; \Omega) = \{(x^*, \alpha) \in X^* \times R :\ \langle x^*, x - \bar{x}\rangle + \alpha\|x - \bar{x}\| \leq 0,\ \forall\, x \in \Omega_T(\bar{x})\}.$$

Clearly,

$$\partial^{lw}_\Omega f(\bar{x}) \supset \partial^w_{\Omega_T(\bar{x})} f(\bar{x}) \supset \partial^w_\Omega f(\bar{x}) \text{ and } N^{lA}(\bar{x}; \Omega) = N^A(\bar{x}; \Omega_T(\bar{x})).$$

In terms of these definitions, the necessary and sufficient conditions of local optimality can be established in the form of (3.1); that is,

$$(0, 0) \in \partial^{lw}_\Omega f(\bar{x}) + N^{lA}(\bar{x}; \Omega). \tag{3.2}$$

Clearly, if $\bar{x}$ is a global optimal solution then it is also a local optimal solution and the optimality condition (3.2) is satisfied if (3.1) holds. Naturally, in a convex case, these conditions coincide.

Classical directional derivative of function f at $\bar{x}$ on direction h is defined as follows:

$$f'(\bar{x}; h) := \lim_{t \downarrow 0} \frac{f(\bar{x} + t\,h) - f(\bar{x})}{t}.$$

We will require the following assumptions hold.

Assumption A:

A1: $f^{'}(\bar{x}; h)$ is defined for all $h \in \mathbf{T}_\Omega(\bar{x})$ and is lower semicontinuous in h;

A2: There exist $\varepsilon, \delta > 0$ such that

$$f(x) - f(\bar{x}) \geq \delta f^{'}(\bar{x}; x - \bar{x}), \ \forall x \in \Omega_T(\bar{x}) \cap B(\bar{x}, \varepsilon). \tag{3.3}$$

The next theorem describes a necessary condition of optimality in the form (3.2) that generalizes Theorem 5 in [8] to any normed spaces by assuming an additional condition $\sigma_L(x^*; \bar{x}) < 1$ and provides the local optimality version of that theorem.

Theorem 3.1. Let $\bar{x} \in \Omega \subset \mathbf{X}$ be a local minimizer of f. Assume that Assumption A holds, there exists $x^* \in \mathbf{U}^*$ such that $\sigma_L(x^*; \bar{x}) < 1$, $\Omega \setminus \{\bar{x}\} \neq \emptyset$ and

$$\bar{\beta} := \inf\{f'(\bar{x}; h) : \ h \in \mathbf{T}_\Omega(\bar{x}) \cap \mathbf{U}\} > 0. \tag{3.4}$$

Then, there exists a nontrivial solution to (3.2); namely, there is $(z^*, \alpha) \in \partial^{lw}_\Omega f(\bar{x})$ such that $(-z^*, -\alpha) \in N^{lA}(\bar{x}; \Omega)$ and $\alpha < \|z^*\|_*$.

Proof. By the assumption, for some $x^* \in U^*$, we have $\sigma_L(x^*; \bar{x}) < 1$. Let $\beta \in (\sigma_L(x^*; \bar{x}), 1)$ and let $\mathbf{C} = \text{cone}\{x \in \mathbf{U} : \ \langle x^*, x\rangle \geq \beta\}$. By Theorem 2.8, cone $\mathbf{C}$ and $\mathbf{T}_\Omega(\bar{x})$ are separable in the sense of Definition 2.7. Therefore by [9, Theorem 4.3], there exists $(y^*, \gamma) \in \partial^w_{\Omega_T(\bar{x})} f(\bar{x}) \subset \partial^{lw}_\Omega f(\bar{x})$ with $y^* \neq 0$ and $\gamma \geq 0$ such that separates the sets $\mathbf{C}$ and $\mathbf{T}_\Omega(\bar{x})$ in the following sense:

$$\langle y^*, y\rangle + \gamma\|y\| < 0 \leq \langle y^*, x\rangle + \gamma\|x\|, \ \forall\, y \in \mathbf{C} \setminus \{0\} \text{ and } \forall\, x \in \mathbf{T}_\Omega(\bar{x}).$$

The rest of proof is the same as in Theorem 4 in [8] to show that there exists $z^* \neq 0$ and $\alpha \geq 0$ such that $(z^*, \alpha) \in \partial^{lw}_\Omega f(\bar{x})$ and

$$\langle z^*, x - \bar{x}\rangle + \alpha\|x - \bar{x}\| \geq 0, \quad \forall x \in \Omega_T(\bar{x}), \tag{3.5}$$

$$\langle z^*, z - \bar{x}\rangle + \alpha\|z - \bar{x}\| < 0, \quad \text{for some } z \notin \Omega_T(\bar{x}). \tag{3.6}$$

Multiplying both sides of (3.5) by -1, we obtain

$$\langle -z^*, x - \bar{x}\rangle - \alpha\|x - \bar{x}\| \leq 0 \ \ \forall x \in \Omega_T(\bar{x}),$$

that means $(-z^*, -\alpha) \in N^{lA}(\bar{x}; \Omega)$. Thus, (3.2) is satisfied.

Now we show that (z^*, α) is a nontrivial solution; that is, $-\alpha > -\| - z^*\|_*$ or $\alpha < \|z^*\|_*$. By contradiction let $\alpha \geq \|z^*\|_*$. Then from the Cauchy-Schwarz inequality it follows that

$$\langle z^*, x - \bar{x}\rangle + \alpha\|x - \bar{x}\| \geq \langle z^*, x - \bar{x}\rangle + \|z^*\|_* \cdot \|x - \bar{x}\| \geq 0, \quad \forall x \in \Omega_T(\bar{x}).$$

This contradicts (3.6). Theorem 3.1 is proved. ■

Theorem 6 from [8] states that condition (3.1) is also a sufficient condition of optimality. This result is straightforward from the definitions of $\partial^w_\Omega f(\bar{x})$ and $N^A(\bar{x}; \Omega)$ and does not require any additional assumptions.

In our case, condition (3.2) may not be a sufficient condition of local optimality if the set Ω is not convex around $\bar{x}$. The next theorem investigates this problem. It shows that if function f is Lipshchitz continuous then the sufficiency of condition (3.2) can be established. Note that this theorem is in any normed spaces, while Theorem 6 from [8] considers the Euclidean norm.

Theorem 3.2. Let (3.2) has a solution, f is Lipshchitz continuous on Ω and

$$\inf\{f^{'}(\bar{x}; h) : \ h \in \ \operatorname{bd}(T_\Omega(\bar{x}))\} \geq \delta > 0. \tag{3.7}$$

Then $\bar{x} \in \Omega$ is a local minimizer of function f on Ω.

Proof. By assumption, there is $(z^*, \alpha) \in \partial^{lw}_\Omega f(\bar{x})$ such that $(-z^*, -\alpha) \in N^{lA}(\bar{x}; \Omega)$. Then there is $\varepsilon > 0$ such that the following holds

$$f(x) - f(\bar{x}) \geq \langle z^*, x - \bar{x}\rangle + \alpha\|x - \bar{x}\| \geq 0; \ \forall x \in \Omega_T(\bar{x}) \cap B(\bar{x}, \varepsilon).$$

We need to show that there is a sufficiently small number $\varepsilon' \leq \varepsilon$ such that this inequality also holds for all $x \in \Omega \cap B(\bar{x}, \varepsilon')$; that is,

$$f(x) - f(\bar{x}) \geq 0; \ \forall x \in \Omega \cap B(\bar{x}, \varepsilon'). \tag{3.8}$$

On the Contrary, assume that (3.8) does not hold. Then for any $n \in N$ satisfying $n\,\varepsilon > 1$ there exists $x_n \in B\left(\bar{x}; \dfrac{1}{n}\right) \cap \Omega$ such that

$$f(x_n) - f(\bar{x}) < 0 \tag{3.9}$$

holds.

Clearly $x_n \notin \Omega_T(\bar{x})$ and $\{x_n - \bar{x}\}_{n\in N}$ approaches 0. Moreover, there exists a convergent subsequence of $z_n := \dfrac{x_n - \bar{x}}{\|x_n - \bar{x}\|}$ as $z_n \in \mathbf{U}$ is bounded. For the sake of simplicity let $z_n \to z$. By definition of $T_\Omega(\bar{x})$ it follows that $z \in T_\Omega(\bar{x})$. On the other hand since $x_n \notin \Omega_T(\bar{x})$ we have $x_n - \bar{x} \notin T_\Omega(\bar{x})$ which means that $z \in \operatorname{bd} T_\Omega(\bar{x}) \cap \mathbf{U}$. By assumption (3.7)

$$f'(\bar{x}; z) \geq \delta > 0. \tag{3.10}$$

Denote $\lambda_n = \|x_n - \bar{x}\|$. Since function f is Lipshchitz, there is $K > 0$ such that for any n the inequality

$$\|f(\bar{x} + \lambda_n z_n) - f(\bar{x} + \lambda_n z)\| \leq \lambda_n K \|z_n - z\|$$

holds and it implies

$$f(\bar{x} + \lambda_n z_n) - f(\bar{x} + \lambda_n z) \geq -\lambda_n K \|z_n - z\|. \tag{3.11}$$

Now, from (3.10) we have

$$f(\bar{x} + \lambda_n z) = f(\bar{x}) + \lambda_n f'(\bar{x}; z) + o(\lambda_n) \geq f(\bar{x}) + \lambda_n \delta + o(\lambda_n);$$

where $\dfrac{o(\lambda_n)}{\lambda_n} \to 0$ as $\lambda_n \to 0$. This together with (3.11) leads to

$$f(\bar{x} + \lambda_n z_n) \geq f(\bar{x}) + \lambda_n \delta + o(\lambda_n) - \lambda_n K \|z_n - z\|$$

or

$$f(\bar{x} + \lambda_n z_n) \geq f(\bar{x}) + \lambda_n \left(\delta + \frac{o(\lambda_n)}{\lambda_n} - K \|z_n - z\|\right).$$

Since $\delta > 0$, $\dfrac{o(\lambda_n)}{\lambda_n} \to 0$ and $\|z_n - z\| \to 0$, we obtain that the inequality $f(\bar{x} + \lambda_n z_n) \geq f(\bar{x})$ holds for sufficiently large n. Taking into account the notations $\lambda_n = \|x_n - \bar{x}\|$ and $z_n = \dfrac{x_n - \bar{x}}{\|x_n - \bar{x}\|}$ we have $\bar{x} + \lambda_n z_n = x_n$ and therefore $f(x_n) \geq f(\bar{x})$ which contradicts (3.9).

Theorem 3.2 is proved. ■

The following example shows that if f is not Lipshchitz then Theorem 3.2 may not be true even if condition (3.7) still holds.

Example 3.3. Let $\mathbf{X} = R^2$ and $\Omega = \{(x_1, x_2) : \ x_1 \geq 0, \ x_2 \leq 2x_1^2\}$. Function f is given by

$$f(x_1, x_2) = \begin{cases} x_1 - x_2, & \text{if } x_2 \leq 0; \\ x_1 - \dfrac{x_2}{x_1}, & \text{if } x_1 \neq 0, \ x_2 \in (0, 2x_1^2); \\ -x_1, & \text{if } x_2 \geq 2x_1^2. \end{cases} \tag{3.12}$$

It is not difficult to observe that f is continuous on Ω. We have $T_\Omega(\bar{x}) = \{(x_1, x_2) : \ x_1 \geq 0, \ x_2 \leq 0\}$ with two boundary directions $\hat{h} = (1, 0)$ and $\tilde{h} = (0, -1)$. The directional derivatives at these directions can be easily calculated to obtain

$$f'(\bar{x}; \hat{h}) = 1 > 0, \quad f'(\bar{x}; \tilde{h}) = 1 > 0.$$

Thus, condition (3.7) holds. We show that function f is not Lipshchitz continuous.

Take any $\varepsilon > 0$ and $h = (0, 1)$. Calculate the directional derivative at point $y^\varepsilon = (\varepsilon, 0)$. We have

$$f'(y^\varepsilon; h) = \lim_{t \downarrow 0\, (t < 2\varepsilon^2)} \frac{f(y^\varepsilon + t\, h) - f(y^\varepsilon)}{t} = -\frac{1}{\varepsilon}.$$

Then, $f'(y^\varepsilon; h) \to -\infty$ as $\varepsilon \downarrow 0$.

Therefore in this example condition (3.7) is satisfied but function f is not Lipshchitz continuous. As a result, Theorem 3.2 is not true; that is, $\bar{x} = (0, 0) \in \Omega$ is a not a local minimizer of f. Indeed, for $x^\varepsilon = (\varepsilon, \varepsilon^2) \in \Omega$ we have $f(x^\varepsilon) = -\varepsilon < 0 = f(\bar{x})$ and $x^\varepsilon \to \bar{x}$ as $\varepsilon \downarrow 0$.

4. Conclusions

In this paper, we introduce local supporting function $\sigma_L(x^*; \bar{x})$ and apply it to characterize non-convex sets at a particular points. The necessary and sufficient conditions of local optimality are derived in terms of three concepts weak subdifferentials, augmented normal cones and the function $\sigma_L(x^*; \bar{x})$. Similar optimality conditions for global optimization are obtained in [8] for Euclidean spaces. This paper generalizes these conditions to local optimization problems and to any finite dimensional normed spaces by applying function $\sigma_L(x^*; \bar{x})$.

References

[1] A. Azimov and R. Gasimov. On weak conjugacy, weak subdifferentials and duality with zero gap in non-convex optimization. *Int. J. Appl. Math*, 1:171–192, 1999.

[2] M. S. Bazaraa and J. J. Goode. On the cones of tangents with applications to mathematical programming. *J. Optimization Theory Appl*, 13:389–426, 1974.

[3] F. H. Clarke. *Optimization and nonsmooth analysis*. Wiley, 1983.

[4] W. Fenchel. Convex cones, sets and functions. Lecture Notes. Princeton University, Princeton, New Jersey, 1951.

[5] A. D. Ioffe. Calculus of Dini subdifferentials of functions and contingent coderivatives of set-valued maps. *Nonlinear Anal., Theory Methods Appl*, 8:517–539, 1984.

[6] R. C. James. Orthogonality and linear functionals in normed linear spaces. *Trans. Am. Math. Soc*, 61:265–292, 1947.

[7] R. Kasimbeyli and M. Mammadov. On weak subdifferentials, directional derivatives, and radial epiderivatives for non-convex functions. *SIAM Journal on Optimization*, 20(2):841–855, 2009.

[8] R. Kasimbeyli and M. Mammadov. Optimality conditions in non-convex optimization via weak subdifferentials. *Nonlinear Analysis: Theory, Methods and Applications*, 74(7):2534–2547, 2011.

[9] R. Kasimbeyli. A nonlinear cone separation theorem and scalarization in non-convex vector optimization. *SIAM J. Optim*, 20(3):1591–1619, 2010.

[10] A. Y. Kruger. On Férechet subdifferentials. *Journal of Mathematical Sciences*, 116(3):3325–3358, 2003.

[11] H. Minkowski. *Theorie der konvexen Körper, insbesondere Begründung ihres Ober flächenbegriffs*. Gesammelte Abhandlungen, II, Teubner, Leipzig. 1911.

[12] B. Mordukhovich and A. Y. Kruger. Necessary optimality conditions in a problem of terminal control with nonfunctional constraints. *[in Russian], Dokl. Akad. Nauk BSSR*, 20(12):1064–1067, 1976.

[13] J. Penot. Calcul sous-différentiel et optimisation. *J. Funct. Anal*, 27:248–276, 1978.

[14] R. T. Rockafellar and R. J. B. Wets. *Variational Analysis*. Grundlehren Der Mathematischen Wissenschaften. Springer, 1998.

[15] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, New Jersey, 1970.

[16] R. T. Rockafellar. Directionally Lipschitzian functions and subdifferential calculus. *Proc. London Math. Soc*, 39:331–355, 1979.

[17] R. T. Rockafellar. *The Theory of Subgradients and Its Applications to Problems of Optimization: Convex and Non-convex Functions*. Helderman Verlag, Berlin, 1981.

[18] W. Rudin. *Real and complex analysis*. McGraw-Hill, Inc, New York, NY, USA, third edition, 1987.

Chapter 4

Necessary and sufficient conditions for global optimality via sigma supporting cone

4.1 Characterizing non-convex sets with conic gap via sigma supporting cone

In this section, the main focus is to characterize nonconvex sets in reflexive Banach spaces. The notion of conic gap for a set is defined and afterwards the sets with conic gap property in reflexive strictly convex space are characterized by applying a new introduced supporting function. The concept of maximal conic gap is defined to measure the size of conic gap of a set. We also use maximal conic gap to characterize nonconvex sets. Then, a global supporting cone based on the supporting function is generalized. A continuous bijective map between supporting cone and a subset of an uniformly convex set is investigated.

Paper:

Characterizing Non-Convex Sets with Conic Gap via Sigma Supporting Cone

Published by:

Proceedings of the 10^{th}IMT-GT ICMSA, 2014

Authors: Sara Hassani[a, b], Musa Mammadov[a, b] and Mina Jamshidi[c]

[a]Federation University Australia, Victoria 3353, Australia

[b]National ICT Australia, VRL, VIC 3010, Australia

[c]Graduate University of advanced technology, Kerman, Iran

Corresponding author:

Sara Hassani

e-mail: Sara.Hassani@nicta.com.au

Characterizing Non-Convex Sets with Conic Gap via Sigma Supporting Cone

Sara Hassania[a,b], Musa Mammadov[a,b] and Mina Jamshidi[c]

[a] *Federation University Australia, Victoria 3353, Australia*
[b] *National Information and Communication Technology Australia*
[c] *Graduate University of advanced technology, Kerman, Iran*
Sara.Hassani@nicta.com.au
m.mammadov@federation.edu.au
m.jamshidi@kgut.ac.ir

Abstract. In this paper, a new supporting function is introduced. Based on this function, the concept of conic gap is defined to characterize non-convex sets. Then sigma normal cone is introduced and its some properties are investigated.

Keywords: Non-convex sets, Conic gap and Sigma supporting cone.

INTRODUCTION

The investigation of geometrical properties of sets is an important problem in many areas of mathematics. The convexity of sets is one of the fundamental notions that are used in many fields. For example, separation theories are based on the convexity of involved sets. Note that convex sets contain the line segment of its any two points.

We briefly mention here an important area in optimization where convexity plays an crucial role. Various kinds of subdifferentials and normal cones have been introduced to describe necessary and sufficient conditions of optimality. They are applicable for different classes of optimization problems, although these concepts (normal cones and subdifferentials) greatly depend on the properties of feasible sets and objective functions. In the convex case, classical notions of subdifferentials and normal cones can be applied; however, for non-convex sets many difficulties arise. Among the many successful approaches developed for non-convex problems, we mention Frechet subdifferentials ([2] and [5]); Clarke's subdifferentials for locally Lipschits functions ([4]); the non-convex limiting Frechet subdifferentials ([2] and [3]) and weak subdifferentials ([1]).

In this paper, we study the structure of non-convex sets. We introduce a new supporting function and then based on this function, we define the concept of conic gap to characterize non-convex sets. These notions can be used to define normal cones for non-convex sets that are important for establishing necessary and sufficient conditions of optimality.

The paper is organized as follows. In the next section, the notions of supporting function and maximal conic gap are introduced. Then these concepts are used to characterize non-convex sets. In the last section, the σ-supporting cone and its properties are presented.

NON-CONVEX SETS WITH MAXIMAL CONIC GAPS

Throughout the paper we assume that X is a reflexive Banach space, unless otherwise stated, with norm $\|.\|$, $\Omega \subseteq X$, $\bar{x} \in \Omega$ and $K = cl\left(cone\ (\Omega - \bar{x})\right)$, where "cl" stands for the closure of a set, and "$cone(A)$" for a given set $A \subseteq X$ stands for

$$cone(A) = \{\lambda x: \lambda \geq 0, x \in A\}.$$

The unit sphere and the unit ball of X are denoted by U and B, respectively:

$$U = \{x \in X: \|x\| = 1\}, \quad B = \{x \in X: \|x\| \leq 1\}.$$

The dual norm of X is denoted by $\|.\|_*$, where $\|.\|_* = \max\{\langle ., x\rangle: x \in U\}$ and $\langle .,. \rangle$ is the scalar product. Note that any continuous linear function attains its supremum on unit ball of reflexive Banach space [9].

The unit sphere and unit ball of dual space of X are denoted by U^* and B^*, respectively. We use "$\rightharpoonup$" to say that a sequence is convergent weakly.

Following [10], we say that Ω has a conic gap at $\bar{x}$ if $K \neq X$.One of the main questions of interest here is to investigate non-convex sets having conic gaps. Therefore, we introduce the following function to characterize the class of such sets.

Let $x^* \in U^*$. We define the function$\sigma_\Omega(x^*; \bar{x})$ for the set Ω at $\bar{x}$:

$$\text{(2.1)} \qquad \sigma_\Omega(x^*; \bar{x}) = Sup_{y \in K \cap U}\langle x^*, y\rangle$$

We present the definition of strictly convex spaces and two propositions used in the rest of paper.

Definition 2.1. ([11], Page 112) Normed space X is called strictly convex if its unit ball is a strictly convex set; i.e., if $x \leq y,\ x, y \in U$and $h = \frac{1}{2}(x + y)$ then $\|h\| < 1$.

The following Proposition from [11, Theorem 5.20] is obtained from Hahn-Banach theorem:

Proposition 2.2. Let$x' \in U$. There exists $x^* \in U^*$ such that

$$\langle x^*, x'\rangle = max_{x \in U}\langle x^*, x\rangle = 1.$$

We will also need the following proposition.

Proposition 2.3. [8] Let X be reflexive strictly convex space and $x^* \in X^*$. Then the maximum of x^*on unit sphere U is unique.

Now we indicate the relation between conic gap and function $\sigma_\Omega(x^*; \bar{x})$ in reflexive strictly convex spaces.

Lemma 2.4. Let X be reflexive strictly convex space. Ω has conic gap at $\bar{x}$ (i.e., $K \neq X$) if and only if

$$\text{(2.2)} \qquad \sigma_\Omega(x^*; \bar{x}) < 1, \text{ for some } x^* \in U^*.$$

Proof: Since $\sigma_\Omega(x^*; \bar{x}) < 1$, we have

$$\sigma_\Omega(x^*; \bar{x}) = \sup_{y \in K \cap U}\langle x^*, y\rangle < 1 = \max_{x \in U}\langle x^*, x\rangle = \langle x^*, x'\rangle, \ \exists x' \in U.$$

Clearly $x' \notin K \cap U$ which means $K \neq X$ and thus Ω has conic gap at $\bar{x}$. Now we show that if $K \neq X$ then $\sigma_\Omega(x^*; \bar{x}) < 1$. Let $x' \in U$ such that $x' \notin K$. By Proposition 2.2, there exists $x^* \in X^*$ such that

$$max_{x \in U}\langle x^*, x\rangle = \langle x^*, x'\rangle = 1.$$

By Proposition 2.3, $x'$is unique, and therefore by definition of $\sigma_\Omega(x^*; \bar{x})$ we have, $\sigma_\Omega(x^*; \bar{x}) < 1$.

Now we investigate conic gaps of non-convex sets; in other words, we define a measure that determines how 'large' is a given conic gap.

Definition 2.5. The maximal conic gap β^* with respect to Ω at $\bar{x}$ is defined as follows:

$$(2.3) \qquad \beta^* := -\inf_{x^* \in U^*}(Sup_{y\in K\cap U}\langle x^*, y\rangle)$$

The minus sign in (2.3), is used to indicate that the value β^* increases with the size of conic gap. It means if $\beta_1^* > \beta_2^*$ then the gap related to the valueβ_1^* is bigger than the other one. Clearly, for any norm the values of β^* are in the interval $[-1,1]$.

First we give an example of two sets with different maximal conic gap values.

Example 2.6. Let the norm be Euclidean, $\Omega_1 = \{(x,y) \in R^2: y \geq 0\}$, $\Omega_2 = \{(x,y) \in R^2: y \geq x \text{ } or \text{ } y \geq -x\}$ and $\bar{x} = (0,0)$. By Definition 2.5, $\beta_1^* = 0$ and $\beta_2^* = \frac{1}{\sqrt{2}}$. As the maximal gap values show, the size of the gap in $\Omega_1$is bigger than Ω_2.

The following result is a characterization of the closed cone cl $(\text{cone}\ (\ \Omega - \bar{x}))$ in terms of β^* in reflexive spaces.

Theorem 2.7. The following hold:

1. If $-1 < \beta^* < 0$, then K is non-convex.
2. If X is a strictly convex space, then $\beta^* = -1$ if and only if $K = X$.
3. If $0 < \beta^* < 1$, then K contains in a half space, i.e. there is $x^* \in X^*$ such that $\langle x^*, x\rangle \leq 0$ for all $x \in K$.

Proof: (1) Assume to the contrary that K is convex. By Hahn-Banach theorem, there exists a linear function x^* such that

$$K \subseteq \{x \in X: \langle x^*, x\rangle \leq 0\}.$$

Then

$$\sigma_\Omega(x^*;\ \bar{x}) = \sup_{y\in K\cap U} \langle x^*, y\rangle \leq 0.$$

which results in $\beta^* \geq 0$. This contradicts the relation $-1 < \beta^* < 0$.

(2) Let $\beta^* = -1$. If $K \neq X$, then by Lemma 2.4, there exists $x^* \in U^*$ such that $\sigma_\Omega(x^*;\ \bar{x}) < 1$, or equivalently $\beta^* < -1$, which is a contradiction.

As any linear functional x^* attains its maximum on unit ball of a reflexive space, so the other side is obvious.

(3) Let $0 < \beta^* < 1$. By the definition of β^*, there exists $x^* \in U^*$ such that

$$\sup_{y\in K\cap U} \langle x^*, y\rangle \leq 0$$

Thus $K \subseteq \{x \in X: \langle x^*, x\rangle \leq 0\}$.

σ – SUPPORTING CONE

In this section, we introduce a new supporting cone which is called “σ-supporting cone” since it is constructed by using function $\sigma_\Omega(x^*;\ \bar{x})$.

Definition 3.1. σ –supporting cone for set Ω at $\bar{x}$ is defined as follows:

$$N^\sigma(\bar{x};\Omega) = cone\ \{x^* \in \boldsymbol{U}^*:\ \sigma_\Omega(x^*;\ \bar{x}) = \text{Sup}_{y\in \boldsymbol{K}\cap\boldsymbol{U}}\langle x^*, y\rangle < 1\}.$$

It is clear that the following representation is true

$$N^{\sigma}(\bar{x};\Omega)=\left\{x^*\in \boldsymbol{X}^*:\ \sigma_{\Omega}(x^*;\ \bar{x})=\operatorname*{Sup}_{y\in \boldsymbol{K}\cap \boldsymbol{U}}\langle x^*,y\rangle<\|x^*\|_*\right\}.$$

First we present the relation between σ-supporting cone and augmented normal cone introduced in ([1] and [10]).

Definition 3.2. Let $\Omega\setminus\bar{x}\neq\emptyset$. The set

$$N^A(\bar{x},\Omega)=\{(x^*,\alpha)\in \boldsymbol{X}^*\times R:\ \langle x^*,x-\bar{x}\rangle+\alpha\ \|x-\bar{x}\|\le 0, for\ all\ x\in\Omega\}$$

is called an augmented normal cone to Ω at $\bar{x}$.

In [10], to derive the necessary and sufficient conditions of optimality, nontrivial elements of augmented normal cone are considered. Note that the element $(x^*,\alpha)\in N^A(\bar{x},\Omega)$is called nontrivial if $\alpha>-\|x^*\|_*$.

Clearly, for any$(x^*,\alpha)\in N^A(\bar{x},\Omega)$, we have

$$\langle x^*,x-\bar{x}\rangle+\alpha\|x-\bar{x}\|\le 0, for\ all\ x\in\Omega$$

and then

$$\alpha\le-\frac{\langle x^*,x-\bar{x}\rangle}{\|x-\bar{x}\|},\qquad for\ all\ \ x\in\Omega, x\neq\bar{x}.$$

This means that if $(x^*,\alpha)\in N^A(\bar{x},\Omega)$ is nontrivial then

$$-\|x^*\|_*<\alpha\le-\mathrm{Sup}_{y\in \boldsymbol{K}\cap \boldsymbol{U}}\langle x^*,y\rangle.$$

Therefore, $(x^*,\alpha)\in N^A(\bar{x},\Omega)$ is nontrivial if and only if

$$\mathrm{Sup}_{y\in \boldsymbol{K}\cap \boldsymbol{U}}\langle x^*,y\rangle<\|x^*\|_*\text{, or } x^*\in N^{\sigma}(\bar{x};\Omega).$$

In other words,

$$x^*\in N^{\sigma}(\bar{x};\Omega)\Leftrightarrow\exists\ \alpha>-\|x^*\|_*\ such\ that\ (x^*,\alpha)\in N^A(\bar{x},\Omega).$$

Now we investigate the relation between $\boldsymbol{X}\setminus\boldsymbol{K}$ and $N^{\sigma}(\bar{x};\Omega)$ in uniformly convex spaces. First we present the definition of uniformly convex spaces.

Definition 3.3. [[6], Section 3.7] $\boldsymbol{X}$ is called uniformly convex if for all$\varepsilon>0$, there exists $\delta>0$ such that for all $\|x\|=\|y\|=1$,

$$\left\|\frac{x+y}{2}\right\|>1-\delta\Leftrightarrow\|x-y\|<\varepsilon.$$

(i.e., if the midpoint of two unit vectors is close to a unit vector, then the two unit vectors are close)

We will use the following proposition in the last lemma.

Proposition 3.4. Let $\boldsymbol{X}$ is uniformly convex, then:

1. $\boldsymbol{X}$ is reflexive [[6], Theorem 3.31 (MilmanPettis)].
2. If $x_j\rightharpoonup x$ weakly and $\|x_j\|\longrightarrow x$, then $x_j\longrightarrow x$ in norm [[6], Proposition 3.32].
3. Unit ball $\boldsymbol{U}$ is weakly sequentially compact [[7], Theorem 7].

It is not difficult to show that $N^{\sigma}(\bar{x};\Omega)=\boldsymbol{X}\setminus\boldsymbol{K}$ in the Euclidean space; however this relation is not true in any normed spaces. The following example in L_1 illustrates that the sets $N^{\sigma}(\bar{x};\Omega)$ and $\boldsymbol{X}\setminus\boldsymbol{K}$ can be very different.

Example 3.5. Let norm of R^2be L_1 (i.e. for any $(x_1, x_2) \in R^2$, $\|(x_1, x_2)\|_1 = |x_1| + |x_2|$). Clearly, the norm of dual space is L_∞ (i.e. for any$(x_1, x_2) \in R^2$, $\|(x_1, x_2)\|_\infty = max\{|x_1|, |x_2|\}$). Let $\Omega = R^2 \setminus \{(x_1, x_2){:}\ x_1 < 0\ and\ x_2 < 0\}$ and $\overline{(x_1, x_2)} = (0,0)$. It is not difficult to observe that there is no $x^* \in \boldsymbol{U}^*$ with $\sigma_\Omega(x^*;\ \bar{x})$<1 $(i.e., N^\sigma(\bar{x}; \Omega) = \emptyset)$ while $\boldsymbol{X} \setminus \boldsymbol{K} = \{(x_1, x_2){:}\ x_1 < 0\ and\ x_2 < 0\}$.

Now we show the relation between $N^\sigma(\bar{x}; \Omega)$ and $\boldsymbol{X} \setminus \boldsymbol{K}$ in uniformly convex spaces.

Lemma 3.6. Let $\boldsymbol{X}$ be a uniformly convex space. Then there is a continuous bijective (one-to-one and onto) mapbetween $N^\sigma(\bar{x}; \Omega)$ and $\boldsymbol{X} \setminus \boldsymbol{K}$.

Proof: Define function ϕ for $x^* \in N^\sigma(\bar{x}; \Omega)$as follows:

$$\phi(x^*) = \|x^*\|_*. x^{\max} \text{ where } x^{\max} = \text{arrgmax}\{\langle x^*, x\rangle{:}\ x \in \mathbf{U}\}.$$

Since $\phi(x^*)$ is positively homogenous, it is enough to show that it is a continuous one-to-one and onto map between $N^\sigma(\bar{x}; \Omega) \cap U^*$ and $(X \setminus K) \cap U$.

For $x^* \in N^\sigma(\bar{x}; \Omega) \cap \boldsymbol{U}^*$we have

$$\sigma_\Omega(x^*;\ \bar{x}) = \text{Sup}_{y \in \boldsymbol{K} \cap \boldsymbol{U}} \langle x^*, y\rangle < 1 = max_{x \in \boldsymbol{U}} \langle x^*, x\rangle = \langle x^*, x^{max}\rangle.$$

From this relation it follows that $x^{max} \notin \boldsymbol{K}$ or $x^{max} \in \boldsymbol{X} \setminus \boldsymbol{K}$. Since $\boldsymbol{X}$ is strictly convex space, $x^{max}$is unique. Thus the function ϕ is well defined on $N^\sigma(\bar{x}; \Omega) \cap \boldsymbol{U}^*$**.**

Let $x' \in (\boldsymbol{X} \setminus \boldsymbol{K}) \cap \boldsymbol{U}$. By Proposition 2.2, there exists $x^* \in \boldsymbol{U}^*$ such that

$$max_{x \in \boldsymbol{U}} \langle x^*, x\rangle = \langle x^*, x'\rangle.$$

As $\boldsymbol{X}$ is strictly convex space, the maximum point x' on $\boldsymbol{U}$is unique. Since $x' \in \boldsymbol{X} \setminus \boldsymbol{K}$ we have $x^* \in N^\sigma(\bar{x}; \Omega)$. Thus ϕ is onto map.

Now we show that $\phi$is continuous. Let $x_k^* \in N^\sigma(\bar{x}; \Omega) \cap \boldsymbol{U}^*$, $x_k^* \to x^*$. We show $\phi(x_k^*)$ approaches $\phi(x^*)$.

Denote$x_k^{max} = \phi(x_k^*)$ and $x^{max} = \phi(x^*)$. Clearly

$$(3.1) \|x_k^*\|_* = max_{x \in \boldsymbol{U}} \langle x_k^*, x\rangle = \langle x_k^*, x_k^{max}\rangle = 1,$$

$$(3.2) \|x^*\|_* = max_{x \in \boldsymbol{U}} \langle x^*, x\rangle = \langle x^*, x^{max}\rangle = 1.$$

Since $\boldsymbol{X}$ is reflexive, there exists a subsequence of $\{x_k^{max}\}_{k \in N}$which is weakly convergent. For the sake of simplicity, let $x_k^{max} \rightharpoonup x'$weakly. Then

$$lim_{k \to \infty}(\langle x_k^*, x_k^{max}\rangle - \langle x^*, x'\rangle) =$$

$$lim_{k \to \infty}(\langle x_k^*, x_k^{max}\rangle - \langle x^*, x_k^{max}\rangle + \langle x^*, x_k^{max}\rangle - \langle x^*, x'\rangle) =$$

$$lim_{k \to \infty}(\langle x_k^* - x^*, x_k^{max}\rangle + \langle x^* x_k^{max} - x'\rangle) = 0.$$

Thus from (3.1) we have $lim_{k \to \infty} \langle x_k^*, x_k^{max}\rangle = \langle x^*, x'\rangle = 1$.

By Proposition [part 3, Proposition 3.4], $x' \in \boldsymbol{U}$ and as $\boldsymbol{X}$ is strictly convex space the maximum point x_k^{max}in (3.2) is unique. Then $x' = x_k^{max}$.

In addition as $\boldsymbol{X}$ is reflexive [part 1, Proposition 3.4], $\{x_k^{max}\}_{k\in N}$ is convergent to x' in norm by [part 2, Proposition 3.4].

Therefore we proved that any weakly convergence subsequence of $\{x_k^{max}\}_{k\in N}$ is norm convergent with the limit point x^{max}.This means that$x_k^{max} \to x^{max}$ or $\phi(x_k^*) \to \phi(x^*)$; that is, ϕ is continuous.

REFERENCES

1. A. Azimov and R. Gasimov, On weak conjugacy, weak subdifferentials and duality with zero gap in non-convex optimization. Int. J. Appl. Math, 1:171-192, 1999.
2. A. Y. Kruger. On Ferechet subdifferentials, Journal of Mathematical Sciences, 116(3):3325-3358, 2003.
3. B. Mordukhovich and A. Y. Kruger, Necessary optimality conditions in a problem of terminal control with nonfunctional constraints. [in Russian], Dokl. Akad. Nauk BSSR, 20(12):1064-1067, 1976.
4. F. H. Clarke. Optimization and nonsmooth analysis. Wiley, 1983.
5. M. S. Bazaraa and J. J. Goode, On the cones of tangents with applications to mathematical programming. J. Optimization Theory Appl, 13:389-426, 1974.
6. H. Brezis, Functional Analysis, Sobolev Spaces and Partial Differential Equations, Springer, 2010.
7. R. C. James, Weak compactness and reflexivity, Israel Journal of Mathematics, 2(2): 101-119, June 1964.
8. R. C. James. Orthogonality and linear functionals in normed linear spaces. Trans. Am. Math. Soc, 61:265-292, 1947.
9. R. C. James. Refexivity and the sup of linear functionals. Ann. of Math, 66:159-169, 1957.
10. R. Kasimbeyli and M. Mammadov, Optimality conditions in nonconvex optimization via weak subdifferentials. Nonlinear Analysis: Theory, Methods and Applications, 74(7):2534-2547, 2011.
11. W. Rudin, Real and complex analysis. McGraw-Hill, Inc, New York, NY, USA, third edition, 1987.

4.2 Sigma supporting cone and optimality conditions in nonconvex problems

In this section, we investigate nonconvex sets that have conic gap. A global supporting function is introduced and afterwards, the notion of sigma supporting cone that is constructed by supporting function, is defined. To determine the size of conic gap, a measure called maximal conic gap is introduced. Finally, the necessary and sufficient conditions for a class of nonconvex and nonsmooth optimization problems in finite dimensional normed spaces by applying weak subdifferentials and global supporting function, for global optimization, is presented.

Paper:

SIGMA SUPPORTING CONE AND OPTIMALITY CONDITIONS IN NON-CONVEX PROBLEMS

Published by:

Far East Journal of Mathematical Sciences (FJMS), 2014

Authors: Sara Hassani[a, b] and Musa Mammadov[a, b]

[a]Federation University Australia, Victoria 3353, Australia

[b]National Information and Communications Technology Australia (NICTA)

Corresponding author:

Sara Hassani

e-mail: sarahassani@students.federation.edu.au

Far East Journal of Mathematical Sciences (FJMS)
© 2014 Pushpa Publishing House, Allahabad, India
Published Online: September 2014
Available online at http://pphmj.com/journals/fjms.htm
Volume 91, Number 2, 2014, Pages 169-190

SIGMA SUPPORTING CONE AND OPTIMALITY CONDITIONS IN NON-CONVEX PROBLEMS

S. Hassani and M. A. Mammadov

National Information and Communications
Technology Australia (NICTA)
Federation University Australia
Victoria 3353, Australia
e-mail: sarahassani@students.federation.edu.au
m.mammadov@federation.edu.au

Abstract

In this paper, a new supporting function for characterizing non-convex sets is introduced. The notions of σ-supporting cone and maximal conic gap are proposed and some properties are investigated. By applying these new notions, we establish the optimality conditions considered in [7] for a broader class of finite dimensional normed spaces in terms of weak subdifferentials.

1. Introduction

The notion of subdifferential plays an important role in optimization theory. It was first introduced as a generalization of the concept of ordinary derivative to deal with optimization problems involving convex and nonsmooth functions by Moreau and Rockafellar [10]. One of the main approaches to derive the necessary conditions of optimality for nonsmooth

Received: April 15, 2014; Accepted: June 26, 2014

2010 Mathematics Subject Classification: 47N10.

Keywords and phrases: non-convex analysis, optimality conditions, supporting cone, non-convex sets.

and non-convex problems is the generalization of the notion of subdifferentials and normal cones.

Nonsmooth and non-convex phenomena have been known for a long time in mathematics and applied sciences. Various kinds of subdifferentials and normal cones have been introduced which are applicable for different classes of optimization problems. We briefly mention here some of major concepts related to this paper. These concepts (normal cones and subdifferentials) in order depend on the properties of the variable space as well as on the objective function. Fréchet subdifferentials were first introduced for finite dimensions in [2] (under the name "lower semidifferentials") (see [8] for some of their properties in an infinite-dimensional setting).

An important generalization of normal cones beyond convexity of functions and sets was done by Clarke [3]. Clarke's directional derivative and its related subdifferential are investigated for the locally Lipschitz functions. Convexity of normal cones and subdifferentials has drawbacks in some cases. To avoid the drawbacks, Mordukhovich and Kruger [9] introduced the non-convex limiting Fréchet normal cone in finite dimensional spaces, and then the concept is extended to infinite dimensional spaces [8].

Augmented normal cones and weak subdifferentials are one of the most useful nonlinear global concepts introduced by Azimov and Gasimov [1]. Recently, Kasimbeyli and Mammadov [6, 7] considered the "necessary" and "sufficient" conditions of optimality for a wide range of non-convex and nonsmooth problems in an Euclidean space. This is the first generalization obtained for non-convex problems in the form of a necessary and sufficient condition.

In this paper, we consider the optimality conditions for non-convex and nonsmooth problems by applying augmented normal cones and weak subdifferentials similar to optimality conditions introduced in [7]. The main purpose is to establish the analogies of these results for a broader class of finite dimensional normed spaces. We introduce three new concepts and then apply them to generalize the necessary and sufficient optimality conditions given in [7].

The paper is organized as follows: a new function for investigating a set with conic gap is established in the next section. σ-supporting cone and its properties are considered in Section 3. In Section 4, we establish the relation of weak subdifferentials with directional derivative. In Section 5, the separation property [5] is established for two specially designed closed cones and then optimality conditions are considered by applying weak subdifferentials.

2. Non-convex Sets with Conic Gaps

Throughout the paper, we assume that $\mathbf{X}$ is a finite dimensional space with norm $\|\cdot\|$, $\Omega \subset \mathbf{X}$, $\bar{x} \in \Omega$ and $\mathbf{K} = \mathrm{cl}(\mathrm{cone}(\Omega - \bar{x}))$, where "cl" stands for the closure of a set, and "$\mathrm{cone}(A)$" for a given set $A \subset \mathbf{X}$ stands for

$$\mathrm{cone}(A) = \{\lambda x : \lambda \geq 0,\ x \in A\}.$$

The unit sphere and the unit ball of $\mathbf{X}$ are denoted by $\mathbf{U}$ and $\mathbf{B}$, respectively: $\mathbf{U} := \{x \in \mathbf{X} : \| x \| = 1\}$, $\mathbf{B} := \{x \in \mathbf{X} : \| x \| \leq 1\}$.

The dual will be denoted by $\mathbf{X}^*$ equipped with the norm $\|\cdot\|_*$, where $\|\cdot\|_* := \max\{\langle \cdot, x\rangle : x \in \mathbf{U}\}$ and $\langle \cdot, \cdot \rangle$ is the scalar product. Note that we use maximum in definition of $\|\cdot\|_*$ as $\mathbf{X}$ is a finite dimensional space. The unit sphere and unit ball of dual space of $\mathbf{X}$ are denoted by $\mathbf{U}^*$ and $\mathbf{B}^*$, respectively.

Following [7], we say that Ω has a *conic gap* at $\bar{x}$ if $\mathbf{K} \neq \mathbf{X}$. One of the questions of interest here is to investigate non-convex sets having conic gaps.

Let $x^* \in \mathbf{U}^*$. We define the following function $\sigma_\Omega(x^*; \bar{x})$ for the set Ω at $\bar{x}$:

$$\sigma_\Omega(x^*; \bar{x}) := \max_{y \in \mathbf{K} \cap \mathbf{U}} \langle x^*, y \rangle. \tag{2.1}$$

Since the set $\mathbf{K} \cap \mathbf{U}$ is closed and bounded and $\mathbf{X}$ is a finite dimensional space, the maximum in (2.1) is attained; that is, function $\sigma_\Omega(x^*; \bar{x})$ is well defined. Moreover, it is not difficult to verify that it is a convex function defined on $\mathbf{X}^*$ and therefore is continuous.

For a convex set Ω, there is another representation for $\sigma_\Omega(x^*; \bar{x})$ given in the following lemma.

Lemma 2.1. *Let Ω be a convex set and $x^* \in \mathbf{U}^*$. Then the following holds*:

$$\sigma_\Omega(x^*; \bar{x}) = \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle x^*, x - \bar{x} \rangle}{\| x - \bar{x} \|}. \tag{2.2}$$

Proof. Clearly, there is a sequence $x_n \in \Omega$, $x_n \to \bar{x}$ such that

$$\xi := \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle x^*, x - \bar{x} \rangle}{\| x - \bar{x} \|} = \lim_{k \to \infty} \frac{\langle x^*, x_k - \bar{x} \rangle}{\| x_k - \bar{x} \|}.$$

Sequence $\left\{ \frac{x_k - \bar{x}}{\| x_k - \bar{x} \|} \right\}_{k \in N}$ is a bounded sequence, so there exists a convergent subsequence of $\left\{ \frac{x_{n_k} - \bar{x}}{\| x_{n_k} - \bar{x} \|} \right\}_{k \in N}$ to x'. As $x_{n_k} \in \Omega$, we have $\frac{x_{n_k} - \bar{x}}{\| x_{n_k} - \bar{x} \|} \in \mathbf{K} \cap \mathbf{U}$ and consequently $x' \in \mathbf{K} \cap \mathbf{U}$. We have

$$\xi = \left\langle x^*, \lim_{k \to \infty} \frac{x_{n_k} - \bar{x}}{\| x_{n_k} - \bar{x} \|} \right\rangle = \langle x^*, x' \rangle \leq \sigma_\Omega(x^*; \bar{x}).$$

Now we show that $\xi \geq \sigma_\Omega(x^*; \bar{x})$. Denote by $\tilde{y} \in \mathbf{K} \cap \mathbf{U}$ a point for which $\langle x^*, \tilde{y}\rangle = \max\limits_{y \in \mathbf{K}\cap\mathbf{U}} \langle x^*, y\rangle$. Then there is a sequence $\tilde{y}_k \in \text{cone}(\Omega - \bar{x}) \cap \mathbf{U}$ such that $\tilde{y}_k \to \tilde{y}$. Moreover, $\tilde{y}_k = \dfrac{\tilde{x}_k - \bar{x}}{\| \tilde{x}_k - \bar{x} \|}$ for some $\tilde{x}_k \in \Omega$. Consider the sequence $z_k = \bar{x} + \frac{1}{k}(\tilde{x}_k - \bar{x})$. Since Ω is convex, we have $z_k \in \Omega$ for all k. Clearly, $z_k \to \bar{x}$ and

$$\xi \geq \lim_{k\to\infty} \frac{\langle x^*, z_k - \bar{x}\rangle}{\| z_k - \bar{x} \|} = \lim_{k\to\infty} \frac{\langle x^*, \tilde{x}_k - \bar{x}\rangle}{\| \tilde{x}_k - \bar{x} \|}$$

$$= \lim_{k\to\infty} \langle x^*, \tilde{y}_k\rangle = \sigma_\Omega(x^*; \bar{x}).$$

Lemma is proved. □

Lemma 2.2. *If* $\sigma_\Omega(x^*; \bar{x}) < 1$ *for some* $x^* \in \mathbf{U}^*$, *then* Ω *has a conic gap at* $\bar{x}$.

Proof. We have

$$\sigma_\Omega(x^*; \bar{x}) = \max_{y \in \mathbf{K}\cap\mathbf{U}} \langle x^*, y\rangle < 1 = \max_{x\in\mathbf{U}} \langle x^*, x\rangle = \langle x^*, x'\rangle, \exists\, x' \in \mathbf{U}.$$

Clearly, $x' \notin \mathbf{K} \cap \mathbf{U}$ which means $\mathbf{K} \neq \mathbf{X}$ and thus Ω has conic gap at $\bar{x}$.

Lemma is proved. □

The following example shows that the inverse of this lemma is not correct. In this example, the set provided has a conic gap at a specific point, however, $\sigma_\Omega(x^*; \bar{x})$ is equal to 1 for all $x^* \in \mathbf{U}^*$.

Example 2.3. Let norm of R^2 be L_1 (i.e., for any $x \in R^2$, $\| x \|_1 = | x_1 | + | x_2 |$). Clearly, the norm of dual space is L_∞ (i.e., for any $x \in R^2$, $\| x \|_\infty = \max\{| x_1 |, | x_2 |\}$). Let $\Omega = R^2 \backslash \{(x, y) \in R^2 : x < 0 \text{ and } y < 0\}$ and

$\bar{x} = (0, 0)$. It is not difficult to observe that there is no $x^* \in \mathbf{U}^*$ with $\sigma_\Omega(x^*; \bar{x}) < 1$; while set Ω has a conic gap at $\bar{x}$ $(\mathbf{K} \neq \mathbf{X})$.

The following lemma explains that the inverse of Lemma 2.2 is true for strictly convex spaces. Before we prove this lemma, we present the definition of strictly convex spaces and two propositions used in the rest of paper.

The following proposition from [11, Theorem 5.20] is obtained from Hahn-Banach theorem:

Proposition 2.4. *Let* $x' \in \mathbf{U}$. *There exists* $x^* \in \mathbf{U}^*$ *such that*

$$\langle x^*, x' \rangle = \max_{x \in \mathrm{U}} \langle x^*, x \rangle = 1.$$

Definition 2.5 [11, p. 112]**.** Normed space $\mathbf{X}$ is called *strictly convex* if its unit ball is a strictly convex set; i.e., if $x \neq y$, $x, y \in \mathbf{U}$ and $h = \frac{1}{2}(x + y)$, then $\| h \| < 1$.

We will also need the following proposition.

Proposition 2.6 [4]**.** *Let* $\mathbf{X}$ *be a strictly convex space and* $x^* \in \mathbf{X}$. *Then the maximum of* x^* *on unit sphere* $\mathbf{U}$ *is unique*.

Now we show that the inverse of Lemma 2.2 is true for strictly convex spaces.

Lemma 2.7. *Let* $\mathbf{X}$ *be a strictly convex space*. *Then* Ω *has conic gap at* $\bar{x}$ *if and only if*

$$\sigma_\Omega(x^*; \bar{x}) < 1 \text{ for some } x^* \in \mathbf{U}^*.$$

Proof. By Lemma 2.2, it is enough to show that if Ω has conic gap at $\bar{x}$ (i.e. $\mathbf{K} \neq \mathbf{X}$) then $\sigma_\Omega(x^*; \bar{x}) < 1$. Let $x' \in \mathbf{U}$ such that $x' \notin \mathbf{K}$. By Proposition 2.4, there exists $x^* \in \mathbf{X}$ such that

$$\max_{x \in \mathbf{U}} \langle x^*, x \rangle = \langle x^*, x' \rangle = 1.$$

By Proposition 2.6, x' is unique, and therefore by the definition of $\sigma_\Omega(x^*; \bar{x})$, we have $\sigma_\Omega(x^*; \bar{x}) < 1$.

Lemma is proved. □

Now we investigate the question of the "size" of a conic gap; in other words, we define a measure that determines how "large" is a given conic gap.

Definition 2.8. The maximal conic gap β^* with respect to Ω at $\bar{x}$ is defined as follows:

$$\beta^* := -\min_{x^* \in \mathbf{U}^*} \sigma_\Omega(x^*; \bar{x}) = -\min_{x^* \in \mathbf{U}^*} \Big(\max_{y \in \mathbf{K} \cap \mathbf{U}} \langle x^*, y \rangle \Big). \tag{2.3}$$

The minus sign in (2.3) is used to indicate that the value β^* increases with the size of conic gap. It means if $\beta_1^* > \beta_2^*$, then the gap related to the value β_1^* is bigger than the other one. Clearly, for any norm, the values of β^* are in the interval $[-1, 1]$.

First, we give an example of two sets with different maximal conic gap values.

Example 2.9. Let the norm be Euclidean, $\Omega_1 = \{(x, y) \in R^2 : y \geq 0\}$, $\Omega_2 = \{(x, y) \in R^2 : y \geq x \text{ or } y \geq -x\}$ and $\bar{x} = (0, 0)$. By Definition 2.8, $\beta_1^* = 0$ and $\beta_2^* = -\frac{1}{\sqrt{2}}$. As the maximal gap values show, the conic gap in Ω_1 is bigger than those in Ω_2.

The following result is a characterization of the closed cone $\mathbf{K} = \operatorname{cl}(\operatorname{cone}(\Omega - \bar{x}))$ in terms of β^*.

Theorem 2.10. *The following hold*:

(1) *If* $-1 < \beta^* < 0$, *then* **K** *is non-convex*.

(2) *If* $\mathbf{X}$ *is a strictly convex space, then* $\beta^* = -1$ *if and only if* $\mathbf{K} = \mathbf{X}$.

(3) *If* $0 \leq \beta^* \leq 1$, *then* $\mathbf{K}$ *contains in a half space, i.e., there is* $x^* \in \mathbf{X}$ *such that* $\langle x^*, x \rangle \leq 0$ *for all* $x \in \mathbf{K}$.

Proof. (1) Assume to the contrary that $\mathbf{K}$ is convex. Clearly, there exists a linear function x^* such that

$$\mathbf{K} \subseteq \{x \in \mathbf{X} : \langle x^*, x \rangle \leq 0\}.$$

Then

$$\sigma_\Omega(x^*; \bar{x}) = \max\{\langle x^*, y \rangle : y \in \mathbf{K} \cap \mathbf{U}\} \leq 0$$

which results in $\beta^* \geq 0.$ This contradicts the relation $-1 < \beta^* < 0.$

(2) Let $\beta^* = -1.$ If $\mathbf{K} \neq \mathbf{X},$ then by Lemma 2.7, there exists $x^* \in \mathbf{U}^*$ such that $\sigma_\Omega(x^*; \bar{x}) < 1,$ or $\beta^* < -1$ which is a contradiction.

The other side is obvious.

(3) Let $0 \leq \beta^* \leq 1.$ By the definition of $\beta^*,$ there exists $x^* \in \mathbf{U}^*$ such that

$$\max\{\langle x^*, y \rangle : y \in \mathbf{K} \cap \mathbf{U}\} \leq 0.$$

Thus, $\mathbf{K} \subset \{x \in \mathbf{X} : \langle x^*, x \rangle \leq 0\}.$

Theorem is proved. □

In Example 2.3 above, the equality $\beta^* = -1$ holds, however $\mathbf{K} \neq \mathbf{X}.$ This shows that the assertion (2) of Theorem 2.10 may not be true for non-strictly convex spaces.

3. σ-supporting Cone

In this section, we introduce a new supporting cone which is called "σ-*supporting cone*" since it is constructed by using function $\sigma_\Omega(x^*; \bar{x}).$

Definition 3.1. σ-supporting cone for set Ω at $\bar{x}$ is defined as follows:

$$N^{\sigma}(\bar{x};\,\Omega) = \operatorname{cone}\{x^{*} \in \mathbf{U}^{*} : \sigma_{\Omega}(x^{*};\,\bar{x}) = \max_{y \in \mathbf{K} \cap \mathbf{U}} \langle x^{*},\, y\rangle < 1\}. \tag{3.1}$$

It is clear that the following representation is true:

$$N^{\sigma}(\bar{x};\,\Omega) = \{x^{*} \in \mathbf{X} : \sigma_{\Omega}(x^{*};\,\bar{x}) = \max_{y \in \mathbf{K} \cap \mathbf{U}} \langle x^{*},\, y\rangle < \| x^{*} \|_{*}\}. \tag{3.2}$$

Clearly, if $N^{\sigma}(\bar{x};\,\Omega)$ is not empty, by continuity of $\sigma_{\Omega}(x^{*};\,\bar{x}),$ then it is an open set.

In Example 2.3 above, there is no $x^{*} \in \mathbf{U}^{*}$ such that $\sigma_{\Omega}(x^{*};\,\bar{x}) < 1$ and consequently $N^{\sigma}(\bar{x};\,\Omega) = \varnothing.$ If $\mathbf{X}$ is a strictly convex space and $\mathbf{X} \backslash \mathbf{K} \neq \varnothing,$ by Lemma 2.7, we have $N^{\sigma}(\bar{x};\,\Omega) \neq \varnothing.$ However, if $\mathbf{X}$ is not strictly convex, then this statement may not be true. In the next two examples, non-strictly convex spaces are considered.

Example 3.2. Let $X = R^{n},$ norm be L_1 and let $(-R^{n}_{+} \backslash \{0_{R^n}\}) \cap \mathbf{K} = \varnothing.$ Then $N^{\sigma}(\bar{x};\,\Omega) \neq \varnothing.$

Indeed, for $y^{*} = (-1,\,-1,\,\ldots,\,-1),$ we have

$$\| y^{*} \|_{*} = \max_{x \in \mathbf{U}} \langle y^{*},\, x\rangle = \max_{x \in \mathbf{U}} (-x_1 - x_2 - \cdots - x_n)$$

$$= \max_{x \in \mathbf{U} \cap (-R^{n}_{+})} (-x_1 - x_2 - \cdots - x_n) > \max_{y \in \mathbf{K} \cap \mathbf{U}} \langle y^{*},\, y\rangle.$$

Thus, $y^{*} \in N^{\sigma}(\bar{x};\,\Omega).$

Example 3.3. Let $X = R^{n},$ norm be L_{∞} and let $y^{*} = \left(-\frac{1}{2},\,-\frac{1}{2},\,\ldots,\,-\frac{1}{2}\right)$ $\notin \mathbf{K}.$ Then $N^{\sigma}(\bar{x};\,\Omega) \neq \varnothing.$

In this case, it can be shown that $y^* \in N^\sigma(\bar{x}; \Omega)$. Indeed, it is clear that for any $x = (x_1, ..., x_n) \in \mathbf{U}$, the inequality $|x_i| \le 1$ holds and therefore

$$\| y^* \|_* = \max_{x \in \mathbf{U}} \langle y^*, x \rangle = \max_{x \in \mathbf{U}} \left(-\frac{1}{2} x_1 - \frac{1}{2} x_2 - \cdots - \frac{1}{2} x_n \right)$$

$$= -\frac{1}{2}(-1) - \cdots - \frac{1}{2}(-1) > \max_{y \in \mathbf{K} \cap \mathbf{U}} \langle y^*, y \rangle$$

which shows that $y^* \in N^\sigma(\bar{x}; \Omega)$.

In the next lemma, a characterization of σ-supporting cone for the Euclidean normed spaces is established. Note that in this case, $\|\cdot\|_* = \|\cdot\|$ and $\mathbf{U}^* = \mathbf{U}$.

Lemma 3.4. *Suppose the norm is Euclidean. Then* $\mathbf{X} \setminus \mathbf{K} = N^\sigma(\bar{x}; \Omega)$.

Proof. If $\mathbf{K} = \mathbf{X}$, then $N^\sigma(\bar{x}; \Omega) = \emptyset$, since $\mathbf{U} \subset \mathbf{K}$. This yields the required equality. Consider the case $\mathbf{K} \neq \mathbf{X}$.

First, we show $\mathbf{X} \setminus \mathbf{K} \subset N^\sigma(\bar{x}; \Omega)$. Let $x^* \in \mathbf{X} \setminus \mathbf{K}$. Since $\mathbf{K} \cap \mathbf{U}$ is a closed set, linear function $\left\langle \frac{x^*}{\| x^* \|}, \cdot \right\rangle$ achieves its maximum on $\mathbf{K} \cap \mathbf{U}$. Let

$$\tilde{y} = \operatorname{argmax} \left\{ \left\langle \frac{x^*}{\| x^* \|}, y \right\rangle : y \in \mathbf{K} \cap \mathbf{U} \right\}.$$

We assumed that norm is Euclidean; therefore, we have

$$\left\langle \frac{x^*}{\| x^* \|}, y \right\rangle \le \left\langle \frac{x^*}{\| x^* \|}, \tilde{y} \right\rangle < 1, \quad \forall y \in \mathbf{K} \cap \mathbf{U}$$

which means

$$\sigma_\Omega \left(\frac{x^*}{\| x^* \|}; \bar{x} \right) = \max_{y \in \mathbf{K} \cap \mathbf{U}} \left\langle \frac{x^*}{\| x^* \|}, y \right\rangle < 1,$$

and consequently $x^* \in N^\sigma(\bar{x}; \Omega)$.

Now we show that $N^{\sigma}(\bar{x};\,\Omega) \subseteq \mathbf{X}\backslash\mathbf{K}$. Let $x^* \in N^{\sigma}(\bar{x};\,\Omega)$ and assume on the contrary that $x^* \in \mathbf{K}$.

Since $\mathbf{K}$ is a cone, $\dfrac{x^*}{\| x^* \|} \in \mathbf{K} \cap \mathbf{U}$ and then $\left\langle x^*, \dfrac{x^*}{\| x^* \|} \right\rangle = \| x^* \|$. This equality leads to a contradiction with $x^* \in N^{\sigma}(\bar{x};\,\Omega)$ in (3.2). This implies $x^* \notin \mathbf{K}$ and $N^{\sigma}(\bar{x};\,\Omega) \subseteq \mathbf{X}\backslash\mathbf{K}$.

Lemma is proved. □

The following example shows that the relation in Lemma 3.4 between set $\mathbf{K}$ and σ-supporting cone may not be true if the norm is not Euclidean.

Example 3.5. Let $\Omega = \{(x,\; y) \in R^2 : |\, y \,| \leq x\}$ and suppose norm of R^2 is L_∞.

For any x^* on the right side of the coordinate system, we have $\max\limits_{y \in \mathbf{K} \cap \mathbf{U}} \langle x^*, y \rangle = \| x^* \|_*$ and then $x^* \notin N^{\sigma}(\bar{x}, \Omega)$ which shows that $N^{\sigma}(\bar{x};\,\Omega)$ $\neq \mathbf{X}\backslash\mathbf{K}$.

Although the relation $\mathbf{X}\backslash\mathbf{K} = N^{\sigma}(\bar{x};\,\Omega)$ is not generally true, there is a one-to-one relation between $\mathbf{X}\backslash\mathbf{K}$ and $N^{\sigma}(\bar{x};\,\Omega)$ when $\mathbf{X}$ is a strictly convex space. The following is a lemma to describe this relation.

Lemma 3.6. *Let* $\mathbf{X}$ *be a strictly convex space. Then there is a continuous bijective* (*one-to-one and onto*) *map between* $N^{\sigma}(\bar{x};\,\Omega)$ *and* $\mathbf{X}\backslash\mathbf{K}$.

Proof. Define a function ϕ for $x^* \in N^{\sigma}(\bar{x};\,\Omega)$ as follows:

$$\phi(x^*) = \| x^* \|_* \cdot x^{\max}, \text{ where } x^{\max} = \operatorname{argmax}\{\langle x^*, x \rangle : x \in \mathbf{U}\rangle.$$

Since $\phi(x^*)$ is positively homogeneous, it is enough to show that it is a continuous one-to-one and onto map between $N^{\sigma}(\bar{x}; \Omega) \cap \mathbf{U}^*$ and $(\mathbf{X} \backslash \mathbf{K}) \cap \mathbf{U}$.

For $x^* \in N^{\sigma}(\bar{x}; \Omega) \cap \mathbf{U}^*$, we have

$$\sigma_{\Omega}(x^*; \bar{x}) = \max_{y \in \mathbf{K} \cap \mathbf{U}} \langle x^*, y \rangle < 1 = \max_{x \in \mathbf{U}} \langle x^*, x \rangle = \langle x^*, x^{\max} \rangle.$$

From this relation, it follows that $x^{\max} \notin \mathbf{K}$ or $x^{\max} \in \mathbf{X} \backslash \mathbf{K}$. Since $\mathbf{X}$ is a strictly convex space, $x^{\max}$ is unique. Thus, the function ϕ is well defined on $N^{\sigma}(\bar{x}; \Omega) \cap \mathbf{U}^*$.

Let $x' \in (\mathbf{X} \backslash \mathbf{K}) \cap \mathbf{U}$. By Proposition 2.4, there exists $x^* \in U^*$ such that

$$\max_{x \in \mathbf{U}} \langle x^*, x \rangle = \langle x^*, x' \rangle.$$

As $\mathbf{X}$ is a strictly convex space, the maximum point x' on $\mathbf{U}$ is unique. Since $x' \in \mathbf{X} \backslash \mathbf{K}$, we have $x^* \in N^{\sigma}(\bar{x}; \Omega)$. Thus, ϕ is an onto map.

Now we show that ϕ is continuous. Let $x_k^* \in N^{\sigma}(\bar{x}; \Omega) \cap \mathbf{U}^*$, $x_k^* \to x^*$. We show $\phi(x_k^*)$ approaches $\phi(x^*)$.

Denote $x_k^{\max} = \phi(x_k^*)$ and $x^{\max} = \phi(x^*)$. Clearly,

$$\| x_k^* \|_* = \max_{x \in \mathbf{U}} \langle x_k^*, x \rangle = \langle x_k^*, x_k^{\max} \rangle = 1, \tag{3.3}$$

$$\| x^* \|_* = \max_{x \in \mathbf{U}} \langle x^*, x \rangle = \langle x^*, x^{\max} \rangle = 1. \tag{3.4}$$

Take any convergent subsequence $x_{k_m}^{\max}$ of $x_k^{\max}$ and let $x_{k_m}^{\max} \to x'$. Then

$$\lim_{m \to \infty} (\langle x_{k_m}^*, x_{k_m}^{\max} \rangle - \langle x^*, x' \rangle) = 0.$$

Thus, from (3.3), we have $\langle x^*, x' \rangle = \lim_{m \to \infty} \langle x_{k_m}^*, x_{k_m}^{\max} \rangle = 1.$

Since $\mathbf{X}$ is a strictly convex space, the maximum point $x^{\max}$ in (3.4) is unique. Then $x' = x^{\max}$.

Therefore, we proved that any convergence subsequence of $\{x_k^{\max}\}_{k \in N}$ has a limit point $x^{\max}$. This means that $x_k^{\max} \to x^{\max}$ or $\phi(x_k^*) \to \phi(x^*)$; that is, ϕ is continuous.

Lemma is proved. □

At the end of this section, we present the relation between σ-supporting cone and augmented normal cone introduced in [7].

Definition 3.7 [7]**.** Let $\Omega \backslash \{\bar{x}\} \neq \varnothing$. The set

$$N^A(\bar{x}; \Omega) := \{(x^*, \alpha) \in X \times R : \langle x^*, x - \bar{x} \rangle + \alpha \| x - \bar{x} \| \leq 0, \forall x \in \Omega\} \tag{3.5}$$

is called an *augmented normal cone* to Ω at $\bar{x}$.

In [7], to derive the necessary and sufficient conditions of optimality, nontrivial elements of augmented normal cone are considered. Note that the element $(x^*, \alpha) \in N^A(\bar{x}; \Omega)$ is called *nontrivial* if $\alpha > -\| x^* \|_*$.

Clearly, for any $(x^*, \alpha) \in N^A(\bar{x}; \Omega)$, we have

$$\langle x^*, x - \bar{x} \rangle + \alpha \| x - \bar{x} \| \leq 0, \quad \forall x \in \Omega,$$

and then

$$\alpha \leq -\frac{\langle x^*, x - \bar{x} \rangle}{\| x - \bar{x} \|}, \quad \forall x \in \Omega, \quad x \neq \bar{x}.$$

This means that if $(x^*, \alpha) \in N^A(\bar{x}; \Omega)$ is nontrivial, then

$$-\| x^* \|_* < \alpha \leq -\max_{y \in \mathbf{K} \cap \mathbf{U}} \langle x^*, y \rangle.$$

This means that

$$x^* \in N^{\sigma}(\bar{x}; \Omega) \Leftrightarrow \exists \alpha > -\| x^* \|_* \text{ such that } (x^*, \alpha) \in N^A(\bar{x}; \Omega);$$

that is, the following lemma is true.

Lemma 3.8. $(x^*, \alpha) \in N^A(\bar{x}; \Omega)$ *is nontrivial if and only if* $x^* \in N^{\sigma}(\bar{x}; \Omega)$.

4. Directional Derivatives and Weak Subdifferentials

In this section, we consider the notion of weak subdifferential, introduced in [1], for any normed spaces. It will be used to establish optimality conditions in the next section. One of the important properties of this notion is its relation with the directional derivatives. This property is established in [7] for the Euclidian norm. In this section, we prove this property for any strictly convex spaces.

Let $f : \Omega \to R$ be a single-valued function. We start with the definition of weak subdifferential.

Definition 4.1. A pair of $(x^*, \alpha) \in X \times R$ is called a *weak subgradient* of *f* at $\bar{x}$ on Ω if

$$f(x) - f(\bar{x}) \geq \langle x^*, x - \bar{x} \rangle + \alpha \| x - \bar{x} \|, \quad \forall x \in \Omega. \tag{4.1}$$

The set

$$\partial_{\Omega}^{w} f(\bar{x}) = \{(x^*, \alpha) \in X \times R : (4.1) \text{ is satisfied}\} \tag{4.2}$$

of all subgradients is called the *weak subdifferential* of *f* at $\bar{x}$ on Ω.

Directional derivative of function *f* at $\bar{x}$ on direction $x - \bar{x}$ is defined as follows:

$$f'(\bar{x}; x - \bar{x}) := \lim_{t \downarrow 0} \frac{f(\bar{x} + t(x - \bar{x})) - f(\bar{x})}{t}.$$

We require that $f'(\bar{x};\, h)$ is defined for all $h \in \mathbf{K} = \mathrm{cl}(\mathrm{cone}(\Omega - \bar{x}))$; accordingly, we will assume that the set Ω satisfies the following condition:

$$\text{For any } h \in \mathbf{K}, \text{ there is } \alpha_h > 0 \text{ such that } \bar{x} + th \in \Omega,\ \forall t \in [0,\, \alpha_h]. \tag{4.3}$$

We will use the following assumption in order to establish some properties of the weak subdifferential and to derive optimality conditions in the next section.

Assumption 1. Suppose that (4.3) holds and f has a directional derivative $f'(\bar{x};\, h)$ at $\bar{x} \in \Omega$ for all $h \in \mathbf{K}$. Moreover, $f'(\bar{x};\, \cdot)$ is lower semicontinuous on $\mathbf{K}$ and there exists $\delta > 0$ such that

$$f(x) - f(\bar{x}) \geq \delta f'(\bar{x};\, x - \bar{x}), \quad \forall x \in \Omega. \tag{4.4}$$

The next theorem is about the relation between weak subdifferentials and directional differentiable functions in strictly convex spaces.

Theorem 4.2. *Let* $\mathbf{X}$ *be a strictly convex space,* $\Omega \subseteq \mathbf{X}$ *and* $\bar{x} \in \Omega$. *Assume that Assumption* 1 *holds and*

$$\inf\{f'(\bar{x};\, h) : h \in \mathbf{K} \cap \mathbf{U}\} > -\infty. \tag{4.5}$$

Then f is weakly subdifferentiable at $\bar{x}$; *that is,* $\partial^w_\Omega f(\bar{x}) \neq \emptyset$. *Moreover, if* $\delta = 1$ *in Assumption* 1, *then*

$$\sup\{\langle x^*,\, h\rangle + \alpha\|h\| : (x^*,\, \alpha) \in \partial^w_\Omega f(\bar{x})\} = f'(\bar{x};\, h), \quad \forall h \in \mathbf{K}. \tag{4.6}$$

Proof. Let $h \in \mathbf{K} \cap \mathbf{U}$. By Proposition 2.4, there exists $x^* \in \mathbf{U}^*$ such that

$$\max_{y \in \mathbf{U}} \langle x^*,\, y\rangle = \langle x^*,\, h\rangle = 1,$$

where h is the unique maximum point according to Proposition 2.6.

Take any $\varepsilon > 0$ and denote $x_1^* = (f'(\bar{x};\, h) - \varepsilon - \alpha_1)x^*$.

We show that there exists sufficiently small α_1 such that $(\delta x_1^*, \delta\alpha_1) \in \partial_\Omega^w f(\bar{x})$.

First, we show that the relation

$$f'(\bar{x};\, z) \geq \langle x_1^*,\, z\rangle + \alpha_1 \| z \|$$

$$= (f'(\bar{x};\, h) - \varepsilon)\langle x^*,\, z\rangle - \alpha_1(\langle x^*,\, z\rangle - 1), \quad \forall z \in \mathbf{K} \cap \mathbf{U} \tag{4.7}$$

is satisfied for some sufficiently small α_1.

Assume to the contrary that this is not true. Then given any sequence $\alpha_n \to -\infty$, there exists $z_n \in \mathbf{K} \cap \mathbf{U}$ such that

$$f'(\bar{x};\, z_n) < (f'(\bar{x};\, h) - \varepsilon)\langle x^*,\, z_n\rangle - \alpha_n(\langle x^*,\, z_n\rangle - 1), \quad \forall n \in N. \tag{4.8}$$

Since $\mathbf{K} \cap \mathbf{U}$ is closed and bounded, there is a convergent subsequence of $\{z_n\}_{n\in N}$. Without loss of generality, assume that z_n converges to $\tilde{z} \in \mathbf{K} \cap \mathbf{U}$.

Let $\tilde{z} \neq h$. As h is a unique maximum point of $\langle x^*, \cdot\rangle$ over the unit ball, the inequality $\langle x^*, \tilde{z}\rangle - 1 < 0$ holds. Then, letting α_n approaches to $-\infty$ in (4.8), we have $f'\langle \bar{x};\, \tilde{z}\rangle = -\infty$, that contradicts (4.5).

Let $\tilde{z} = h$ and consequently $\langle x^*, \tilde{z}\rangle - 1 = 0$. Then by taking limit in (4.8) and using lower semicontinuity of directional derivative $f'(\bar{x};\, \cdot)$, as well as the inequality $\langle x^*,\, z_n\rangle - 1 \leq 0,\ \forall n,$ we have

$$f'(\bar{x};\, h) \leq \liminf_{n\to-\infty} f'(\bar{x};\, z_n) \leq (f'(\bar{x};\, h) - \varepsilon)\langle x^*,\, h\rangle = f'(\bar{x};\, h) - \varepsilon.$$

Since $\varepsilon > 0$, this is again a contradiction.

Therefore, (4.7) holds for some sufficiently small α_1. Take any $x \in \Omega$, $x \neq \bar{x}$. Then $\dfrac{x - \bar{x}}{\| x - \bar{x} \|} \in \mathbf{K} \cap \mathbf{U}$ and from (4.7), we obtain

$$f'(\bar{x};\, x - \bar{x}) \geq \langle x_1^*,\, x - \bar{x} \rangle + \alpha_1 \| x - \bar{x} \|, \quad \forall x \in \Omega,\ x \neq \bar{x}.$$

This relation also holds for $x = \bar{x}$. Then, from (4.4), it follows that

$$f(x) - f(\bar{x}) \geq \delta f'(\bar{x};\, x - \bar{x}) \geq \langle \delta x_1^*,\, x - \bar{x} \rangle + \delta\alpha_1 \| x - \bar{x} \|, \quad \forall x \in \Omega.$$

Thus, $(\delta x_1^*,\, \delta\alpha_1) \in \partial_\Omega^w f(\bar{x})$; that is, the set of weak subdifferentials is not empty.

Now consider the case $\delta = 1$. Then from $(x_1^*,\, \alpha_1) \in \partial_\Omega^w f(\bar{x})$, we have

$$\sup\{\langle x^*,\, h \rangle + \alpha \| h \| : (x^*,\, \alpha) \in \partial_\Omega^w f(\bar{x})\}$$

$$\geq \langle x_1^*,\, h \rangle + \alpha_1 \| h \|$$

$$= (f'(\bar{x};\, h) - \varepsilon)\langle x^*,\, h \rangle - \alpha_1(\langle x^*,\, h \rangle - 1) = f'(\bar{x};\, h) - \varepsilon.$$

Since this relation holds for any $\varepsilon > 0$, we obtain

$$\sup\{\langle x^*,\, h \rangle + \alpha \| h \| : (x^*,\, \alpha) \in \partial_\Omega^w f(\bar{x})\} \geq f'(\bar{x};\, h). \tag{4.9}$$

On the other hand, it is not difficult to show that for any $(x^*,\, \alpha) \in \partial_\Omega^w f(\bar{x})$, the inequality

$$f'(\bar{x};\, h) \geq \langle x^*,\, h \rangle + \alpha \| h \|,$$

and consequently

$$f'(\bar{x};\, h) \geq \sup\{\langle x^*,\, h \rangle + \alpha \| h \| : (x^*,\, \alpha) \in \partial_\Omega^w f(\bar{x})\} \tag{4.10}$$

holds. Then, for given $h \in \mathbf{K} \cap \mathbf{U}$, the required relation (4.6) follows from (4.9) and (4.10). Since both sides in (4.6) are superlinear in *h*, it is also true for all $h \in \mathbf{K}$.

Theorem is proved. □

5. Weak Subdifferentials and Optimality Condition

In this section, we consider the necessary and sufficient conditions for a class of non-convex and nonsmooth optimization problems in finite dimensional normed spaces by applying weak subdifferentials, augmented normal cones and the function $\sigma_{\Omega}(x^*; \bar{x})$. Similar optimality conditions are considered in [7] for the Euclidean space.

We will use the following, so-called, the separation property introduced in [5].

Definition 5.1 [5]**.** Let **C** and **K** be closed cones of a normed space **X**. Let $\widetilde{\mathbf{C}}$ and $\widetilde{K}^{\partial}$ be the closure of the sets $\mathrm{co}(\mathbf{C} \cap \mathbf{U})$ and $\mathrm{co}((\mathrm{bd}(\mathbf{K}) \cap \mathbf{U}) \cup \{0_{\mathbf{X}}\})$. The cones **C** and **K** are said to have the *separation property* with respect to the norm $\|\cdot\|$ if

$$\widetilde{\mathbf{C}} \cap \widetilde{K}^{\partial} = \varnothing. \tag{5.1}$$

Take any positive number $\beta < 1$ and $x^* \in \mathbf{U}^*$. Consider the cone

$$\mathrm{C} = \mathrm{cone}\{x \in \mathrm{U} : \langle x^*, x\rangle \geq \beta\}. \tag{5.2}$$

In the following theorem, we show that under some conditions on the σ-supporting cone defined in Section 3, the cones **C** and **K** satisfy the separation property.

Theorem 5.2. *Let there exists* $x^* \in \mathbf{U}^*$ *such that* $\sigma_{\Omega}(x^*; \bar{x}) < 1$. *Then given any positive number* $\beta \in (\sigma_{\Omega}(x^*; \bar{x}), 1)$, *cones* **C** *and* **K** *satisfy the separation property*.

Proof. By the assumption of the theorem,

$$\max_{y \in \mathbf{K} \cap \mathbf{U}} \langle x^*, y\rangle = \sigma_{\Omega}(x^*; \bar{x}) < 1 = \| x^* \|_* = \max_{x \in \mathbf{U}} \langle x^*, x\rangle. \tag{5.3}$$

Denote $\alpha = \sigma_{\Omega}(x^*; \bar{x})$ and take any $\beta > 0$ such that

$$\alpha = \max_{y \in \mathbf{K} \cap \mathbf{U}} \langle x^*, y \rangle = \sigma_{\Omega}(x^*; \bar{x}) < \beta < 1. \tag{5.4}$$

Since $\mathbf{U}$ is closed, there exists $a \in \mathbf{U}$ such that $\langle x^*, a \rangle = \| x^* \|_* = 1.$ Then $a \in \mathbf{C}$ and $\mathbf{C} \neq \varnothing.$

Denote $\tilde{\mathbf{C}} = \mathrm{cl}(\mathrm{co}(\mathbf{C} \cap \mathbf{U}))$ and $\tilde{K}^{\partial} = \mathrm{cl}(\mathrm{co}((\mathrm{bd}(\mathbf{K}) \cap \mathbf{U}) \cup \{0_{\mathbf{X}}\})).$ We need to prove $\tilde{\mathbf{C}} \cap \tilde{K}^{\partial} = \varnothing.$

First, we show that for any $x \in \tilde{\mathbf{C}},$ the inequality $\langle x^*, x \rangle \geq \beta$ holds. Let $x \in \mathrm{co}(\mathbf{C} \cap \mathbf{U}).$ Then the following representation is true $x = \sum_{i=1}^{n+1} \alpha_i x_i;$ where $x_i \in \mathbf{C} \cap \mathbf{U}$ and $\sum_{i=1}^{n+1} \alpha_i = 1.$ As $x_i \in \mathbf{C} \cap \mathbf{U},$ from (5.2), we have

$$\langle x^*, x \rangle = \sum_{i=1}^{n+1} \alpha_i \langle x^*, x_i \rangle \geq \beta. \tag{5.5}$$

From continuity of x^* and (5.5), for any $x \in \mathrm{cl}(\mathrm{co}(\mathbf{C} \cap \mathbf{U})),$ we have $\langle x^*, x \rangle \geq \beta.$

It is clear from (5.4) that for any $y \in \mathbf{K} \cap \mathbf{U},$ the relation $\langle x^*, y \rangle \leq \alpha < \beta$ holds. Since $\beta > 0,$ we have $\langle x^*, 0 \rangle = 0 < \beta.$ Thus, $\langle x^*, y \rangle \leq \max\{\alpha, 0\} < \beta$ for any $y \in \tilde{K}^{\partial}.$ Therefore, $\tilde{\mathbf{C}} \cap \tilde{K}^{\partial} = \varnothing.$

Theorem is proved. □

The next theorem describes the necessary condition of optimality that generalizes theorem 4 in [7] for any normed spaces under some conditions to the function $\sigma_{\Omega}(x^*; \bar{x}).$

Theorem 5.3. *Let* $\Omega \subset \mathbf{X}$ *and* $f : \Omega \to R$ *be a given function. Assume that* $\bar{x}$ *is a minimizer of f over* Ω *and there exists* $x^* \in U^*$ *such that*

$\sigma_\Omega(x^*; \bar{x}) < 1.$ *Let* $\Omega \setminus \{\bar{x}\} \neq \emptyset,$ *Assumption* 1 *holds and*

$$\overline{\beta} := \inf\{f'(\bar{x}; h) : h \in \mathbf{K} \cap \mathbf{U}\} > 0. \tag{5.6}$$

Then there exists $(z^*, \alpha) \in \partial_\Omega^w f(\bar{x})$ *with* $z^* \neq 0,\ \alpha \geq 0$ *such that*

$$\langle z^*, x - \bar{x} \rangle + \alpha \| x - \bar{x} \| \geq 0, \quad \forall x \in \Omega, \tag{5.7}$$

$$\langle z^*, z - \bar{x} \rangle + \alpha \| z - \bar{x} \| < 0 \ \textit{for some} \ z \notin \Omega. \tag{5.8}$$

Proof. Let $\sigma_\Omega(x^*; \bar{x}) < 1$ for $x^* \in U^*$. By Theorem 5.2, there exists cone $\mathbf{C}$ such that $\mathbf{C}$ and $\mathbf{K}$ are separable in the sense of Definition 5.1. Therefore, by [5, Theorem 4.3], there exists $(y^*, \gamma) \in \partial_\Omega^w f(\bar{x})$ with $y^* \neq 0$ and $\gamma \geq 0$ such that it separates the sets $\mathbf{C}$ and $\mathbf{K}$ in the following sense:

$$\langle y^*, y \rangle + \gamma \| y \| < 0 \leq \langle y^*, x \rangle + \gamma \| x \|, \quad \forall y \in \mathbf{C} \setminus \{0\} \text{ and } \forall x \in \mathbf{K}.$$

The rest of proof is the same as in Theorem 4 in [7].

Theorem is proved. □

The following theorem is about the existence of nontrivial solutions to

$$(0, 0) \in \partial_\Omega^w f(\bar{x}) + N^A(\bar{x}; \Omega). \tag{5.9}$$

Theorem 5.4. *Let all the conditions of Theorem* 5.3 *hold. Then there exists a nontrivial solution to* (5.9).

Proof. All the conditions of Theorem 5.3 hold. Therefore, there exists $(z^*, \alpha) \in \partial_\Omega^w f(\bar{x})$ such that $z^* \neq 0,\ \alpha \geq 0$ and (5.7), (5.8) hold. Multiplying both sides of (5.7) by $-1,$ we obtain

$$\langle -z^*, x - \bar{x} \rangle - \alpha \| x - \bar{x} \| \leq 0, \quad \forall x \in \Omega,$$

that means $(-z^*, -\alpha) \in N^A(\bar{x}; \Omega).$ Thus, (5.9) is satisfied.

4.3 Optimality conditions via weak subdifferentials in Reflexive Banach spaces

In this section, the setting under consideration is reflexive Banach spaces. The directional derivative of a function in reflexive strictly convex space, is presented as supremum of weak subgradients. Based on supporting cone introduced in the previous sections, a separation property for two cones is investigated. We consider necessary and sufficient optimality conditions by applying two notions of weak subdifferential and supporting function.

Paper:

Optimality condition via weak subdifferentials in reflexive Banach spaces

Published by:

Turkish Journal of Mathematics, 2015

Authors: S. Hassani[a, b], M. A. Mammadov[a, b] and M. Jamshidi[c]

[a]Federation University of Austalia, Victoria 3353, Australia

[b]National Information and Communications Technology Australia (NICTA)

[c]Graduate University of advanced technology, Kerman, Iran

Corresponding author:

Sara Hassani

e-mail: Sara.Hassani@nicta.com.au

Turkish Journal of Mathematics

http://journals.tubitak.gov.tr/math/

Research Article

Turk J Math
(2017) 41: 1 – 8
© TÜBİTAK
doi:10.3906/mat-1505-38

Optimality conditions via weak subdifferentials in reflexive Banach spaces

Sara HASSANI[1,2,*], **Musa MAMMADOV**[1,2], **Mina JAMSHIDI**[3]
[1]Federation University of Australia, Ballarat, Victoria, Australia
[2]National Information and Communications Technology Australia, Sydney, Australia
[3]Faculty of Science and New Technologies, Kerman Graduate University of Advanced Technology, Kerman, Iran

Received: 11.05.2015 • **Accepted/Published Online:** 29.02.2016 • **Final Version:** 16.01.2017

Abstract: In this paper the relation between the weak subdifferentials and the directional derivatives, as well as optimality conditions for nonconvex optimization problems in reflexive Banach spaces, are investigated. It partly generalizes several related results obtained for finite dimensional spaces.

Key words: Supporting cone, weak subdifferential and nonconvex optimization

1. Introduction

The generalization of concepts of ordinary derivatives and normal cones plays an important role in the study of necessary and sufficient conditions of optimality for nonsmooth and nonconvex optimization problems. The notion of subdifferentials was introduced by Rockafellar [17] to deal with optimization problems involving convex and nonsmooth functions. Since then, different notions of subdifferentials and normal cones have been introduced, which are applicable for different classes of optimization problems. We mention here the concepts of the Fréchet subdifferential [3, 15], Clarke's subdifferential [4], and limiting Fréchet subdifferentials [15, 16].

In [1, 2], the notion of a supporting cone was introduced and led to so-called weak subdifferentials. To eliminate the duality gap in nonconvex programming, an augmented Lagrangian is used that is constructed by supporting cones [2, 5, 6]. Later in [12], the concept of an augmented dual cone was introduced in Banach spaces and a special class of sublinear functions was defined by using the elements of the augmented dual cone; it was shown that two closed cones possessing a separation property can be separated by using a zero sublevel set of some function from this class. Recently, these concepts were used in [13, 14] to obtain necessary and sufficient conditions of optimality for a wide range of nonconvex and nonsmooth problems in Euclidean space.

In this paper, we study optimality conditions for nonconvex nonsmooth problems in reflexive Banach spaces by applying augmented normal cones and weak subdifferentials. The main purpose is to establish the analogies of the main results obtained in [14] for infinite dimensional normed spaces by using the notion of the supporting cone introduced in [8].

The paper is organized as follows. The main notations, definitions, and preliminaries are presented in the next section. In Section 3, we establish the relation of weak subdifferentials with the directional derivatives in reflexive Banach spaces. Optimality conditions in infinite dimensional normed spaces by applying weak subdifferentials are presented in section 4.

*Correspondence: sara.hassani@nicta.com.au

2. Notations

Throughout the paper we assume that $\mathbf{X}$ is a reflexive Banach space with norm $\|\cdot\|$ unless otherwise stated. Let $\Omega \subset \mathbf{X}$ and $\bar{x} \in \Omega$. We will use the notation $\mathbf{K} = \mathrm{cl}\,(\mathrm{cone}(\Omega - \bar{x}))$ where "cl" stands for the closure of a set, and "cone(A)" for a given set $A \subset \mathbf{X}$ stands for

$$\mathrm{cone}(A) = \{\lambda x : \ \lambda \geq 0, \ x \in A\}.$$

The unit sphere and the unit ball of $\mathbf{X}$ are denoted by $\mathbf{U}$ and $\mathbf{B}$, respectively:

$$\mathbf{U} := \{x \in \mathbf{X} : \ \|x\| = 1\}, \ \ \mathbf{B} := \{x \in \mathbf{X} : \ \|x\| \leq 1\}.$$

The dual norm of $\mathbf{X}$ is denoted by $\|\cdot\|_*$, where $\|\cdot\|_* := \max\{\langle\cdot, x\rangle : \ x \in \mathbf{U}\}$ where $\langle\cdot,\cdot\rangle$ is the scalar product (note that any continuous linear function attains its supremum on a unit ball of reflexive Banach space [10]). The unit sphere and unit ball of dual space of $\mathbf{X}$ are denoted by $\mathbf{U}^*$ and $\mathbf{B}^*$, respectively.

We say that Ω has a conic gap at $\bar{x}$ if $\mathbf{K} \neq \mathbf{X}$ (this property for set Ω at $\bar{x}$ was called "cone-shaped" in [14]). In [7, 8] a new supporting function was introduced to characterize the class of nonconvex sets having conic gaps. Given $x^* \in \mathbf{U}^*$, this supporting function $\sigma_\Omega(x^*; \bar{x})$ for the set Ω at $\bar{x}$ is defined as:

$$\sigma_\Omega(x^*; \bar{x}) := \sup_{y \in \mathbf{K} \cap \mathbf{U}} \langle x^*, y\rangle. \tag{2.1}$$

We present the definition of strictly convex spaces and three propositions used in the remainder of this paper.

Definition 2.1 *(page 112, [18]) Normed space $\mathbf{X}$ is called strictly convex if its unit ball is a strictly convex set, i.e. if $x \neq y$, $x, y \in \mathbf{U}$, and $h = \frac{1}{2}(x+y)$ then $\|h\| < 1$.*

The following proposition is an extension of the Hahn–Banach theorem [18, Theorem 5.20]:

Proposition 2.2 *Let $x' \in \mathbf{U}$. There exists $x^* \in \mathbf{U}^*$ such that*

$$\langle x^*, x'\rangle = \max_{x \in \mathbf{U}} \langle x^*, x\rangle = 1.$$

Proposition 2.3 *[9] Let $\mathbf{X}$ be reflexive strictly convex space and $x^* \in \mathbf{X}^*$. Then the maximum of x^* on unit sphere $\mathbf{U}$ is unique.*

Proposition 2.4 *[11, Theorem 7] Unit ball $\mathbf{U}$ of reflexive space is weakly sequentially compact.*

A new supporting cone, a "*$\sigma-$supporting cone*", constructed by using function $\sigma_\Omega(x^*; \bar{x})$ is introduced in the next definition.

Definition 2.5 *A σ-supporting cone for the set Ω at $\bar{x}$ is defined as follows:*

$$N^\sigma(\bar{x}; \Omega) = cone\{x^* \in \mathbf{U}^* : \ \sigma_\Omega(x^*; \bar{x}) = \sup_{y \in \mathbf{K} \cap \mathbf{U}} \langle x^*, y\rangle < 1\}. \tag{2.2}$$

We show that the following representation is true for a σ-supporting cone:

$$N^{\sigma}(\bar{x};\Omega) = \{x^* \in \mathbf{X}^* : \ \sigma_{\Omega}(x^*;\bar{x}) = \sup_{y\in \mathbf{K}\cap \mathbf{U}} \langle x^*, y\rangle < \|x^*\|_*\}. \tag{2.3}$$

Denote

$$\mathbf{C} = \text{cone}\{x^* \in \mathbf{U}^* : \ \sigma_{\Omega}(x^*;\bar{x}) = \sup_{y\in \mathbf{K}\cap \mathbf{U}} \langle x^*, y\rangle < 1\};$$

$$\mathbf{D} = \{x^* \in \mathbf{X}^* : \ \sigma_{\Omega}(x^*;\bar{x}) = \sup_{y\in \mathbf{K}\cap \mathbf{U}} \langle x^*, y\rangle < \|x^*\|_*\},$$

and let $x^* \in \mathbf{C}$. Clearly, we have $\frac{x^*}{\|x^*\|_*} \in \mathbf{U}^*$ and therefore

$$\sigma_{\Omega}(\frac{x^*}{\|x^*\|_*};\bar{x}) = \sup_{y\in \mathbf{K}\cap \mathbf{U}} \langle \frac{x^*}{\|x^*\|_*}, y\rangle = \frac{1}{\|x^*\|_*} \sup_{y\in \mathbf{K}\cap \mathbf{U}} \langle x^*, y\rangle < 1. \tag{2.4}$$

From (2.4), we obtain $\sup_{y\in \mathbf{K}\cap \mathbf{U}} \langle x^*, y\rangle < \|x^*\|_*$ and that means $x^* \in \mathbf{D}$.

Thus, $\mathbf{C} \subset \mathbf{D}$. To show the inverse inclusion, take any $x^* \in \mathbf{D}$; that is,

$$\sup_{y\in \mathbf{K}\cap \mathbf{U}} \langle x^*, y\rangle < \|x^*\|_*. \tag{2.5}$$

Dividing both sides of equation (2.5) by $\|x^*\|_*$, we obtain

$$\sup_{y\in \mathbf{K}\cap \mathbf{U}} \langle \frac{x^*}{\|x^*\|_*}, y\rangle < 1,$$

and that means $\frac{x^*}{\|x^*\|_*} \in \{x^* \in \mathbf{U}^* : \ \sigma_{\Omega}(x^*;\bar{x}) = \sup_{y\in \mathbf{K}\cap \mathbf{U}} \langle x^*, y\rangle < 1\}$ and consequently $x^* \in \mathbf{C}$ and $\mathbf{D} \subset \mathbf{C}$.

3. Directional derivatives and weak subdifferentials

The notion of a weak subdifferential, introduced in [1] for any normed spaces, will be used to establish optimality conditions in the next section. One of the important properties of this notion is its relation with the directional derivatives. This property was established in [14] for the Euclidean norm. In this section we prove this property for any reflexive Banach space that is strictly convex.

Let $f : \Omega \to R$ be a single-valued function. We start with the definition of weak subdifferential.

Definition 3.1 *A pair of* $(x^*, \alpha) \in \mathbf{X}^* \times R$ *is called a weak subgradient of* f *at* $\bar{x}$ *on* Ω *if*

$$f(x) - f(\bar{x}) \geq \langle x^*, x - \bar{x}\rangle + \alpha\|x - \bar{x}\|, \quad \forall x \in \Omega. \tag{3.1}$$

The set

$$\partial^w_{\Omega} f(\bar{x}) = \{(x^*, \alpha) \in \mathbf{X}^* \times R : (3.1) \text{ is satisfied}\} \tag{3.2}$$

of all subgradients is called the weak subdifferential of f at $\bar{x}$ on Ω.

The directional derivative of function f at $\bar{x}$ on direction $x - \bar{x}$ is defined as follows:

$$f^{'}(\bar{x}; x - \bar{x}) := \lim_{t\downarrow 0} \frac{f(\bar{x} + t(x - \bar{x})) - f(\bar{x})}{t}.$$

We will use the following assumption in order to establish some properties of the weak subdifferential and to derive optimality condition.

Assumption 1. *Suppose that* $\mathbf{K} = cone(\Omega - \bar{x})$ *is a closed set and* f *has a directional derivative* $f^{'}(\bar{x}; h)$ *at* $\bar{x} \in \Omega$ *for all* $h \in \mathbf{K}$. *Moreover,* $f^{'}(\bar{x}; \cdot)$ *is lower semicontinuous on* $\mathbf{K}$ *and there exists* $\delta > 0$ *such that*

$$f(x) - f(\bar{x}) \geq \delta f^{'}(\bar{x}; x - \bar{x}), \; \forall x \in \Omega. \tag{3.3}$$

The next theorem is about the relation between weak subdifferentials and directionally differentiable functions in reflexive Banach spaces that are strictly convex.

Theorem 3.2 *Let* $\mathbf{X}$ *be a reflexive strictly convex space,* $\Omega \subseteq \mathbf{X}$ *and* $\bar{x} \in \Omega$. *Assume that Assumption 1 holds and*

$$inf\{f^{'}(\bar{x}; h) : \; h \in \mathbf{K} \cap \mathbf{U}\} > -\infty. \tag{3.4}$$

Then f *is weakly subdifferentiable at* $\bar{x}$; *that is,* $\partial^w_\Omega f(\bar{x}) \neq \emptyset$. *Moreover, if* $\delta = 1$ *in Assumption 1, then*

$$\sup\{\langle x^*, h\rangle + \alpha\|h\| : \; (x^*, \alpha) \in \partial^w_\Omega f(\bar{x})\} = f^{'}(\bar{x}; h), \;\; \forall \, h \in \mathbf{K}. \tag{3.5}$$

Proof: Let $h \in \mathbf{K} \cap \mathbf{U}$. By Proposition 2.2, there exists $x^* \in \mathbf{U}^*$ such that

$$\max_{y \in \mathbf{U}} \langle x^*, y\rangle = \langle x^*, h\rangle = 1,$$

where h is the unique maximum point according to Proposition 2.3.

Take any $\epsilon > 0$ and denote $x_1^* = (f^{'}(\bar{x}; h) - \epsilon - \alpha_1)\, x^*$ where $\alpha_1 \in (-\infty, \infty)$.

We show that there exists sufficiently small α_1 such that $(\delta\, x_1^*, \delta\, \alpha_1) \in \partial^w_\Omega f(\bar{x})$.

First we show that the relation

$$f^{'}(\bar{x}; z) \geq \langle x_1^*, z\rangle + \alpha_1\|z\| = (f^{'}(\bar{x}; h) - \epsilon)\langle x^*, z\rangle - \alpha_1(\langle x^*, z\rangle - 1), \;\; \forall z \in \mathbf{K} \cap \mathbf{U}, \tag{3.6}$$

is satisfied for some sufficiently small α_1.

Assume to the contrary that this is not true. Then given any sequence $\alpha_n \to -\infty$, there exists $z_n \in \mathbf{K} \cap \mathbf{U}$ such that

$$f^{'}(\bar{x}; z_n) < (f^{'}(\bar{x}; h) - \epsilon)\langle x^*, z_n\rangle - \alpha_n(\langle x^*, z_n\rangle - 1), \;\; \forall \, n \in N. \tag{3.7}$$

By Proposition 2.4, there is a weakly convergent subsequence of $\{z_n\}_{n \in N}$. Without loss of generality assume that z_n converges weakly to $\tilde{z} \in \mathbf{U}$.

Let $\tilde{z} \neq h$. As h is a unique maximum point of $\langle x^*, \cdot\rangle$ over the unit ball, the inequality $\langle x^*, \tilde{z}\rangle - 1 < 0$ holds. Then, letting α_n approach to $-\infty$ in (3.7), we have $f^{'}(\bar{x}; \tilde{z}) = -\infty$, which contradicts (3.4).

Let $\tilde{z} = h$ and consequently $\langle x^*, \tilde{z}\rangle - 1 = 0$. Then by taking the limit in (3.7) and using the lower semicontinuity of the directional derivative $f^{'}(\bar{x}; \cdot)$, as well as the inequality $\langle x^*, z_n\rangle - 1 \leq 0, \; \forall n$, we have

$$f^{'}(\bar{x}; h) \leq \liminf_{n \to -\infty} f^{'}(\bar{x}; z_n) \leq (f^{'}(\bar{x}; h) - \epsilon)\langle x^*, h\rangle = f^{'}(\bar{x}; h) - \epsilon.$$

Since $\epsilon > 0$, this is again a contradiction.

Therefore, (3.6) holds for some sufficiently small α_1. Take any $x \in \Omega$, $x \neq \bar{x}$. Then $\frac{x-\bar{x}}{\|x-\bar{x}\|} \in \mathbf{K} \cap \mathbf{U}$ and from (3.6) we obtain

$$f^{'}(\bar{x}; x-\bar{x}) \geq \langle x_1^*, x-\bar{x}\rangle + \alpha_1\|x-\bar{x}\|, \quad \forall x \in \Omega, x \neq \bar{x}.$$

This relation also holds for $x = \bar{x}$. Then from (3.3) it follows that

$$f(x) - f(\bar{x}) \geq \delta\, f^{'}(\bar{x}; x-\bar{x}) \geq \langle \delta\, x_1^*, x-\bar{x}\rangle + \delta\,\alpha_1\|x-\bar{x}\|, \quad \forall x \in \Omega.$$

Thus, $(\delta\, x_1^*, \delta\,\alpha_1) \in \partial^w_\Omega f(\bar{x})$; that is, the set of weak subdifferentials is not empty.

Now consider the case $\delta = 1$. From $(x_1^*, \alpha_1) \in \partial^w_\Omega f(\bar{x})$, we have

$$\sup\{\langle x^*, h\rangle + \alpha\|h\| : \ (x^*, \alpha) \in \partial^w_\Omega f(\bar{x})\} \geq \langle x_1^*, h\rangle + \alpha_1\|h\| =$$

$$(f^{'}(\bar{x}; h) - \epsilon)\langle x^*, h\rangle - \alpha_1(\langle x^*, h\rangle - 1) = f^{'}(\bar{x}; h) - \epsilon.$$

Since this relation holds for any $\epsilon > 0$, we obtain

$$\sup\{\langle x^*, h\rangle + \alpha\|h\| : \ (x^*, \alpha) \in \partial^w_\Omega f(\bar{x})\} \geq f^{'}(\bar{x}; h). \tag{3.8}$$

On the other hand, it is not difficult to show that, for any $(x^*, \alpha) \in \partial^w_\Omega f(\bar{x})$, the inequality

$$f^{'}(\bar{x}; h) \geq \langle x^*, h\rangle + \alpha\|h\|$$

and consequently

$$f^{'}(\bar{x}; h) \geq \sup\{\langle x^*, h\rangle + \alpha\|h\| : \ (x^*, \alpha) \in \partial^w_\Omega f(\bar{x})\} \tag{3.9}$$

hold. Then, for given $h \in \mathbf{K} \cap \mathbf{U}$, the required relation (3.5) follows from (3.8) and (3.9). Since both sides in (3.5) are superlinear in h, it is also true for all $h \in \mathbf{K}$.

□

4. Weak subdifferentials and optimality condition

In this section we consider the necessary and sufficient conditions for a class of nonconvex and nonsmooth optimization problems in reflexive Banach spaces by applying weak subdifferentials, augmented normal cones, and the function $\sigma_\Omega(x^*; \bar{x})$. Similar optimality conditions are considered in [14] and [7] for the Euclidean space and any finite normed space, respectively.

We will use the following so-called separation property introduced in [12].

Definition 4.1 *[12] Let $\mathbf{C}$ and $\mathbf{K}$ be closed cones of a normed space $\mathbf{X}$. Let $\tilde{\mathbf{C}}$ and $\tilde{K}^\partial$ be the closure of the sets co$(\mathbf{C} \cap \mathbf{U})$ and co$((bd(\mathbf{K}) \cap \mathbf{U}) \cup \{0_\mathbf{X}\})$, respectively. The cones $\mathbf{C}$ and $\mathbf{K}$ are said to have the separation property with respect to the norm $\|\cdot\|$ if*

$$\tilde{\mathbf{C}} \cap \tilde{K}^\partial = \emptyset. \tag{4.1}$$

Take any positive number $\beta < 1$ and $x^* \in \mathbf{U}^*$. Consider the cone

$$\mathbf{C} = \operatorname{cone}\{x \in \mathbf{U} : \langle x^*, x\rangle \geq \beta\}. \tag{4.2}$$

In the following theorem we show that under some conditions on the σ-supporting cone, the cones $\mathbf{C}$ and $\mathbf{K}$ satisfy the separation property.

Theorem 4.2 *Let $\mathbf{X}$ be a reflexive Banach space and let there exist $x^* \in \mathbf{U}^*$ such that $\sigma_\Omega(x^*;\bar{x}) < 1$. Then given any positive number $\beta \in (\sigma_\Omega(x^*;\bar{x}), 1)$, cones $\mathbf{C}$ and $\mathbf{K}$ satisfy the separation property.*

Proof: By the assumption of the theorem

$$\sup_{y\in\mathbf{K}\cap\mathbf{U}} \langle x^*, y\rangle = \sigma_\Omega(x^*;\bar{x}) < 1 = \|x^*\|_* = \max_{x\in\mathbf{U}}\langle x^*, x\rangle. \tag{4.3}$$

Denote $\alpha = \sigma_\Omega(x^*;\bar{x})$ and take any $\beta > 0$ such that

$$\alpha = \sup_{y\in\mathbf{K}\cap\mathbf{U}} \langle x^*, y\rangle = \sigma_\Omega(x^*;\bar{x}) < \beta < 1. \tag{4.4}$$

Since $\mathbf{X}$ is reflexive, there exists $a \in \mathbf{U}$ such that $\langle x^*, a\rangle = \|x^*\|_* = 1$. Then $a \in \mathbf{C}$ and $\mathbf{C} \neq \emptyset$.

Denote $\tilde{\mathbf{C}} = \operatorname{cl}(\operatorname{co}(\mathbf{C}\cap\mathbf{U}))$ and $\tilde{K}^\partial = \operatorname{cl}(\operatorname{co}((\operatorname{bd}(\mathbf{K})\cap\mathbf{U})\cup\{0_\mathbf{X}\}))$. We need to prove $\tilde{\mathbf{C}}\cap\tilde{K}^\partial = \emptyset$.

First we show that for any $x \in \tilde{\mathbf{C}}$ the inequality $\langle x^*, x\rangle \geq \beta$ holds. Let $x \in \operatorname{co}(\mathbf{C}\cap\mathbf{U})$. Then the following representation is true $x = \sum_{i=1}^{n(x)} \alpha_i\tilde{x}_i$ for some $n(x) \in \mathbf{N}$, where $\tilde{x}_i \in \mathbf{C}\cap\mathbf{U}$ and $\sum_{i=1}^{n(x)}\alpha_i = 1$. As $\tilde{x}_i \in \mathbf{C}\cap\mathbf{U}$, from (4.2) we have

$$\langle x^*, x\rangle = \sum_{i=1}^{n(x)} \alpha_i\langle x^*, \tilde{x}_i\rangle \geq \beta. \tag{4.5}$$

Let $x \in \operatorname{cl}(\operatorname{co}(\mathbf{C}\cap\mathbf{U}))$, which means there exists sequence x_n in $\mathbf{C}\cap\mathbf{U}$ such that x_n is convergent to x weakly and consequently $\langle x^*, x_n\rangle \to \langle x^*, x\rangle$. Then by (4.5), we have $\langle x^*, x\rangle \geq \beta$.

It is clear from (4.4) that for any $y \in \mathbf{K}\cap\mathbf{U}$, the relation $\langle x^*, y\rangle \leq \alpha < \beta$ holds. Since $\beta > 0$, we have $\langle x^*, 0\rangle = 0 < \beta$. Thus, $\langle x^*, y\rangle \leq \max\{\alpha, 0\} < \beta$ for any $y \in \tilde{K}^\partial$. Therefore, $\tilde{\mathbf{C}}\cap\tilde{K}^\partial = \emptyset$. □

The condition of reflexivity of $\mathbf{X}$ is important in the proof of Theorem 4.2, although it is our opinion that it can be relaxed. We provide an example below where the space $\mathbf{X}$ is not reflexive but the separation property is still valid.

Example 4.3 *Consider the Banach space $\mathbf{X} = C^0([0,1], R)$ with the norm $\|f\|_\infty = \max\{f(x) : x \in [0,1]\}$. Clearly $\mathbf{X}$ is not reflexive [4].*

Let the linear continuous function x^* be defined as follows:

$$\langle x^*, f\rangle := \int_0^{\frac{1}{2}} f(t)dt - \int_{\frac{1}{2}}^1 f(t)dt \text{ where } f \in \mathbf{X}.$$

We show that $x^* \in \mathbf{U}^*$. Clearly, $\langle x^*, f\rangle \leq 1$ for any $f \in \mathbf{U}$ and hence $\|x^*\|_* \leq 1$. Consider a sequence of functions $f_n(x)$ defined by

$$f_n(x) = \begin{cases} 1 & \text{if} \quad x \in [0, \frac{1}{2} - \frac{1}{n}] \\ -nx + \frac{n}{2} & \text{if} \quad x \in [\frac{1}{2} - \frac{1}{n}, \frac{1}{2} + \frac{1}{n}] \\ -1 & \text{if} \quad x \in [\frac{1}{2} + \frac{1}{n}, 1]. \end{cases}$$

It is easy to check that $\|f_n\|_\infty = 1$ *and* $\langle x^*, f_n\rangle = 1 - \frac{2}{n}$*. Therefore,* $\|x^*\|_* \geq \sup_n(1-\frac{2}{n}) = 1$*; that is,* $\|x^*\|_* \in \mathbf{U}^*$*.*

Now we consider any set Ω *satisfying* $\sigma_\Omega(x^*;\bar{x}) < 1$*. Take an arbitrary number* $\beta \in (\sigma_\Omega(x^*;\bar{x}), 1)$*. First we show that* $\mathbf{C} \neq \emptyset$*.*

Clearly there is n *such that* $\langle x^*, f_n\rangle = 1-2/n \in (\beta, 1)$*; hence,* $\mathbf{C} \neq \emptyset$*. Now we show that for any* $x \in \tilde{\mathbf{C}}$*, the inequality* $\langle x^*, x\rangle \geq \beta$ *holds. By following the proof of Theorem 4.2, for any* $x \in co(\mathbf{C}\cap\mathbf{U})$*, the inequality* $\langle x^*, x\rangle \geq \beta$ *holds. Let* $x \in cl(co(\mathbf{C}\cap\mathbf{U}))$*, which means there exists sequence* x_n *in* $co(\mathbf{C}\cap\mathbf{U}))$ *such that* x_n *is convergent to* x*, i.e.* $\|x_n - x\| \to 0$*. We have*

$$\langle x^*, x_n - x\rangle \leq \|x_n - x\| \to 0;$$

that means $\langle x^*, x_n\rangle \to \langle x^*, x\rangle$ *and consequently,* $\langle x^*, x\rangle \geq \beta$*. Hence, for any* $x \in \tilde{\mathbf{C}}$ *the inequality* $\langle x^*, x\rangle \geq \beta$ *holds.*

Again, by following the proof of Theorem 4.2, it is easy to show that $\langle x^*, y\rangle \leq \max\{\alpha, 0\} < \beta$ *for any* $y \in \tilde{K}^\partial$*. Therefore,* $\tilde{K}^\partial \cap \tilde{\mathbf{C}} = \emptyset$*.*

□

The next theorem describes the necessary condition of optimality that generalizes Theorem 4 in [14] to any reflexive spaces by applying the function $\sigma_\Omega(x^*;\bar{x})$.

Theorem 4.4 *Let* $\mathbf{X}$ *be a reflexive Banach space,* $\Omega \subset \mathbf{X}$ *and* $f : \Omega \to R$ *be a given function. Assume that* $\bar{x}$ *is a minimizer of* f *over* Ω *and there exists* $x^* \in U^*$ *such that* $\sigma_\Omega(x^*;\bar{x}) < 1$*. Letting* $\Omega\setminus\{\bar{x}\} \neq \emptyset$*, Assumption 1 holds and*

$$\bar{\beta} := \inf\{f^{'}(\bar{x}; h) : \ h \in \mathbf{K}\cap\mathbf{U}\} > 0. \tag{4.6}$$

Then there exists $(z^*, \alpha) \in \partial^w_\Omega f(\bar{x})$ *with* $z^* \neq 0$*,* $\alpha \geq 0$ *such that*

$$\langle z^*, x - \bar{x}\rangle + \alpha\|x - \bar{x}\| \geq 0, \ \ \forall x \in \Omega, \tag{4.7}$$

$$\langle z^*, z - \bar{x}\rangle + \alpha\|z - \bar{x}\| < 0, \ \ \textit{for some } z \notin \Omega. \tag{4.8}$$

Proof: Let $\sigma_\Omega(x^*;\bar{x}) < 1$ for $x^* \in U^*$. By Theorem 4.2, there exists cone $\mathbf{C}$ such that $\mathbf{C}$ and $\mathbf{K}$ are separable in the sense of Definition 4.1. Therefore, by [12, Theorem 4.3], there exists $(y^*, \gamma) \in \partial^w_\Omega f(\bar{x})$ with $y^* \neq 0$ and $\gamma \geq 0$ such that it separates the sets $\mathbf{C}$ and $\mathbf{K}$ in the following sense:

$$\langle y^*, y\rangle + \gamma\|y\| < 0 \leq \langle y^*, x\rangle + \gamma\|x\|, \ \forall\, y \in \mathbf{C}\setminus\{0\} \text{ and } \forall\, x \in \mathbf{K}.$$

The rest of the proof is the same as in Theorem 4 in [14].

□

The following theorem presents sufficient conditions guaranteeing the existence of nontrivial solutions to

$$(0, 0) \in \partial^w_\Omega f(\bar{x}) + N^A(\bar{x}; \Omega). \tag{4.9}$$

Theorem 4.5 *Let all the conditions of Theorem 4.4 hold. Then there exists a nontrivial solution to 4.9.*

Proof: All conditions of Theorem 4.4 hold. Therefore, there exists $(z^*, \alpha) \in \partial^w_\Omega f(\bar{x})$ such that $z^* \neq 0$, $\alpha \geq 0$ and (4.7),(4.8) hold. Multiplying both sides of (4.7) by -1, we obtain

$$\langle -z^*, x - \bar{x}\rangle - \alpha\|x - \bar{x}\| \leq 0 \ \ \forall x \in \Omega,$$

which means $(-z^*, -\alpha) \in N^A(\bar{x}; \Omega)$. Thus, (4.9) is satisfied.

Now we show that (z^*, α) is a nontrivial solution; that is, $-\alpha > -\|-z^*\|_*$ or $\alpha < \|z^*\|_*$. By contradiction let $\alpha \geq \|z^*\|_*$. Then from the Cauchy–Schwarz inequality it follows that

$$\langle z^*, x - \bar{x}\rangle + \alpha\|x - \bar{x}\| \geq \langle z^*, x - \bar{x}\rangle + \|z^*\|_* \cdot \|x - \bar{x}\| \geq 0, \ \ \forall x.$$

This contradicts (4.8). $\square$

References

[1] Azimov AY, Gasimov RN. On weak conjugacy, weak subdifferentials and duality with zero gap in nonconvex optimization. Int J Appl Math 1999; 1: 171-192.

[2] Azimov AY, Gasimov RN. Stability and duality of nonconvex problems via augmented Lagrangian. Cybern Syst Anal 2002; 38: 412-421.

[3] Bazaraa MS, Goode JJ. On the cones of tangents with applications to mathematical programming. J Opt Theo Appl 1974; 13: 389-426.

[4] Clarke FH. Optimization and Nonsmooth Analysis. New York, NY, USA: John Wiley, 1983.

[5] Gasimov RN. Augmented Lagrangian duality and nondifferentiable optimization methods in nonconvex programming. J Global Optim 2002; 24: 187-203.

[6] Gasimov RN, Rubinov AM. On augmented Lagrangians for optimization problems with a single constraint. J Global Optim 2004; 28: 153-173.

[7] Hassani S, Mammadov M. Sigam supporting cone and optimality conditions in non-convex problems. Far East Journal of Mathematical Sciences 2014; 91: 169-190.

[8] Hassani S, Mammadov M, Jamshidi M. Characterizing non-convex sets with conic gap via the elements of sigma supporting cone. In: Proceedings of the 10thIMT-GT ICMSA2014, Kuala Terengganu, Malaysia, 2014, pp. 143-148.

[9] James RC. Orthogonality and linear functionals in normed linear spaces. T Am Math Soc 1947; 61: 265-292.

[10] James RC. Reflexivity and the sup of linear functionals. Ann Math 1957; 66: 159-169.

[11] James RC. Weak compactness and reflexivity. Israel J Math 1964; 2: 101-119.

[12] Kasimbeyli R. A nonlinear cone separation theorem and scalarization in nonconvex vector optimization. SIAM J Optimiz 2010; 20:1591-1619.

[13] Kasimbeyli R, Mammadov M. On weak subdifferentials, directional derivatives, and radial epiderivatives for non-convex functions. SIAM J Optimiz 2009; 20: 841-855.

[14] Kasimbeyli R, Mammadov M. Optimality conditions in nonconvex optimization via weak subdifferentials. Nonlinear Anal-Theor 2011; 74: 2534-2547.

[15] Kruger AY. On Férechet subdifferentials. Journal of Mathematical Sciences 2003; 116: 3325-3358.

[16] Mordukhovich BS, Kruger AY. Necessary optimality conditions in a problem of terminal control with nonfunctional constraints. Dokl Akad Nauk BSSR 1976; 20: 1064-1067.

[17] Rockafellar RT. Convex Analysis. Princeton, NJ, USA: Princeton University Press, 1970.

[18] Rudin W. Real and Complex Analysis. New York, NY, USA: McGraw-Hill, 1987.

Chapter 5

Optimality conditions for difference of topical functions and infinite horizon optimization

5.1 Characterizations of minimal elements of topical functions on semimodules with applications

In this section, we consider optimization problems where objective function is the difference of two topical functions. The extended valued topical functions, in this part, are defined on a semimodule with values in a semifield. We first give characterizations of the superdifferential of topical functions and then we characterize minimal elements of the upper support set of extended valued topical functions. Finally, as an application, we present a necessary and sufficient condition for global maximum of the difference of two strictly topical functions.

Paper:

Characterizations of Minimal Elements of Topical Functions on Semimodules with Applications

Published by:

Numerical Functional Analysis and Optimization, 2016

Authors: S. Hassani[b, c] and H. Mohebi[a]

[a]Shahid Bahonar university of Kerman, Iran

[b]Federation University Australia, Victoria 3353, Australia

[c]National Information and Communications Technology Australia (NICTA)

Corresponding author:

Sara Hassani

e-mail: sarahassani@students.federation.edu.au

e-mail: hmohebi@mail.uk.ac.ir

Linear Algebra and its Applications 520 (2017) 104–124

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Characterizations of minimal elements of topical functions on semimodules with applications

Sara Hassani [a], Hossein Mohebi [b,*,1]

[a] *Faculty of Science and Technology, Federation University Australia, VIC, 3353, Australia*
[b] *Department of Mathematics, Shahid Bahonar University of Kerman, P.O. Box 76169133, 7616914111, Kerman, Iran*

A R T I C L E I N F O

Article history:
Received 26 July 2016
Accepted 15 January 2017
Available online 17 January 2017
Submitted by B. Lemmens

MSC:
06F25
06F30
26A48
90C46

Keywords:
Semifield
Semimodule
Minimal element
Topical function
Support set
b-Complete
Superdifferential
DC-functions
Global optimization

A B S T R A C T

In this paper, we first give characterizations of the superdifferential of extended valued topical functions defined on a semimodule with values in a semifield. Next, we characterize minimal elements of the upper support set of extended valued topical functions. Finally, as an application, we present a necessary and sufficient condition for global maximum of the difference of two strictly topical functions defined on a semimodule.

* Corresponding author.
E-mail addresses: Sara.Hassani@nicta.com.au (S. Hassani), hmohebi@uk.ac.ir, hossein.mohebi@unsw.edu.au (H. Mohebi).
[1] Present address: Department of Applied Mathematics, University of New South Wales, Sydney, NSW, 2052, Australia.

http://dx.doi.org/10.1016/j.laa.2017.01.014
0024-3795/

1. Introduction

Topical functions have arisen in several contexts, and the term “topical function” is due to Gunawardena and Keane [13]. Topical functions are intensively studied (see [8,11,12] and the references therein) and they have many applications in various parts of applied mathematics, in particular, in the modelling of discrete event systems (see [11,12]).

Topical functions are also interesting from a different point of view, namely as a tool in the study of the so-called downward sets. Downward sets arise in the study of some problems of mathematical economics and game theory and also in the study of inequality systems involving increasing functions (see [21]).

One of the most important global optimization problems, is that of minimizing a DC-functions (difference of two convex functions), that is,

$$\text{minimize } h(x) \text{ subject to } x \in X,$$

where $h := g - f$ and f, g are convex functions. In a general case, DC-functions can be replaced by DAC-functions (difference of two abstract convex functions). In particular, minimizing of the difference of two increasing and co-radiant (ICR) functions [6], and also, minimizing of the difference of two increasing co-radiant and quasi-concave functions (see; for example, [7,18]).

Recently, topical functions $f : X \longrightarrow K$ and related classes of functions have studied in [4,10,17,19,21,25,26,22,24], where X is a b-complete idempotent semimodule over a b-complete idempotent semifield K. We recall that a function $f : X \longrightarrow K$ is called topical if it is increasing (i.e., the relations $x', x'' \in X$, $x' \leq x''$ imply $f(x') \leq f(x'')$, where $\leq$ denotes the canonical order on X, respectively on K, defined by $x \leq y$ if and only if $x \oplus y = y$ for all $x \in X$ and all $y \in X$, respectively by $\lambda \leq \mu$ if and only if $\lambda \oplus \mu = \mu$ for all $\lambda \in K$ and all $\mu \in K$), and homogeneous (i.e., $f(\lambda x) = \lambda f(x)$ for all $x \in X$ and all $\lambda \in K$, where $\lambda x := \lambda \otimes x$ and $\lambda f(x) := \lambda \otimes f(x)$; the fact that we use the same notations for addition $\oplus$ both in X and in K and for multiplication $\otimes$ both in $K \times X$ and in K will lead to no confusion).

Extended valued topical functions with values in $\overline{K} := K \cup \{\top\}$, where $\top := \sup K$ (possible does not belong to K), have been investigated in [2,25,26]. In fact, by defining new multiplications $\dot{\otimes}$ and $\otimes$ on $\overline{K}$ and introducing two classes of elementary functions, abstract convexity and abstract concavity of extended valued topical functions have been presented. In this paper, we first give characterizations of the superdifferential of this class of functions. Next, we characterize minimal elements of the upper support set of extended valued topical functions. Finally, as an application, we present a necessary and sufficient condition for global maximum of the difference of two topical functions defined on a semimodule.

The paper has the following structure: In Section 2, we provide some preliminary definitions and results relative to semimodules, semifields and topical functions. Characterizations of the superdifferential of extended valued topical functions are given in Section 3. In Section 4, we first characterize minimal elements of the upper support set of extended valued topical functions. Finally, as an application, we present a necessary and sufficient condition for global maximum of the difference of two strictly topical functions defined on a semimodule. Section 5, includes with a discussion on conclusions.

2. Preliminaries

Let $(K, \oplus, \otimes, \varepsilon, e)$ be a semiring with idempotent addition, where the idempotency of $\oplus$ means that $a \oplus a = a$ for all $a \in K$. The addition $\oplus$ and multiplication $\otimes$ have the neutral elements ε and e, respectively. We recall (see [24]) that the idempotent addition $\oplus$ defines an order relation $\leq$ on the semiring K: $a \leq b \Leftrightarrow a \oplus b = b$ for all $a, b \in K$ (with the convention $a \otimes b = ab$). We say $\lambda < \lambda'$, if $\lambda \leq \lambda'$ and $\lambda \neq \lambda'$, where $\lambda, \lambda' \in K$. The notations $\vee$, $\wedge$ are the lattice operations supremum and infimum on K, respectively, which will be defined with respect to this order relation.

We recall (see [25,26,24]) the following definitions.

Let $(X, \oplus', \otimes')$ be a semimodule over a semiring K with idempotent addition $\oplus'$, where the idempotency of $\oplus'$ means $x \oplus' x = x$ for all $x \in X$, and $\oplus' : X \times X \mapsto X$ is defined by $\oplus'(x, y) = x \oplus' y$ for all $x, y \in X$, and $\otimes' : K \times X \mapsto X$ is defined by $\otimes'(\lambda, x) = \lambda \otimes' x = \lambda x$ for all $x \in X$ and all $\lambda \in K$ (with the convention $\lambda \otimes' x = \lambda x$). The idempotent addition $\oplus'$ defines an order relation $\leq$ on the semimodule X (over the semiring K): $x \leq y \Leftrightarrow x \oplus' y = y$ for all $x, y \in X$. We denote the addition on K and X with the same notation $\oplus$. Similarly, the notations $\vee$, $\wedge$ are the lattice operations supremum and infimum on X, respectively, which will be defined with respect to the above order relation.

We recall (see [24]) that a function $f : X \to K$ is called topical, if f has the following properties:

(1) f is increasing, i.e., $x, y \in X, x \leq y \Longrightarrow f(x) \leq f(y)$.
(2) f is homogeneous, i.e., $f(\lambda x) = \lambda f(x)$ for all $\lambda \in K$ and all $x \in X$.

An example (see [24]) of a topical function is $\vee$-linear function on X with values in K, i.e., the function which is homogeneous and satisfying the following condition:

$$f(x \vee y) = f(x) \vee f(y), \quad \forall\, x, y \in X.$$

In particular, an example (see [24]) of a topical function is max-linear function on $\mathbb{R}^n$, i.e.,

(a) $f(\lambda \mathbb{1}_n + x) = \lambda + f(x) \;\; \forall\, \lambda \in \mathbb{R}\,, \forall\, x \in \mathbb{R}^n$,
(b) $f(x \vee y) = \max\{f(x), f(y)\} \;\; \forall\, x, y \in \mathbb{R}^n$, where $\mathbb{1}_n := (1, 1, \ldots, 1) \in \mathbb{R}^n$ and $\lambda \mathbb{1}_n + x = (\lambda + x_1, ..., \lambda + x_n)$, (for more details see also [26]).

The following definitions are well-known.

A function $f : X \longrightarrow K$ is called abstract concave with respect to a set H of K-valued functions defined on X, or H-concave, if there exists a set $U \subseteq H$ such that $f(x) = \inf_{\ell \in U} \ell(x)$ for each $x \in X$ (see [20]).

We recall (see [14,15,25,26,24]) that a semiring $(K, \oplus, \otimes)$ or a semimodule $(X, \oplus, \otimes)$ (over a semiring K) which is closed under the sum $\oplus$ of any subset (order-) bounded from above and the multiplication $\otimes$ distributes over such sums is called boundedly complete (b-complete).

We recall (see [9,25,26,24]) that a commutative semiring K in which every $\mu \in K\backslash\{\varepsilon\}$ is invertible for multiplication $\otimes$ is called semifield.

Throughout the paper we use the notation $\perp$ for $\inf X$, that is, $\perp := \inf X$.

Assumption (A0). For each $x \in X$ and $y \in X \setminus \{\perp\}$, the set $\{\lambda \in K : x \leq \lambda y\}$ is non-empty.

Throughout the paper we assume that X is a b-complete idempotent semimodule over a b-complete idempotent semifield K, and the supremum of each (order-) bounded from above subset of K belongs to K.

It is worth noting that if $(K, \oplus, \otimes)$ is a b-complete semifield with idempotent addition $\oplus$, then the infimum of each non-empty subset of K belongs to K (see [5,16]).

In the sequel, we accept without any special mention that K has at least two elements, and hence $e \neq \varepsilon$.

We say that a semifield K has the property $(\mathcal{C})$ if

$$\inf\{\lambda : \lambda \in K,\ \lambda > \varepsilon\} = \varepsilon.$$

In fact, we need a continuity property for the semifield K similar to the field of real numbers, or a continuous lattice. But, the property $(\mathcal{C})$ is a particular case of the continuity property of a continuous lattice, and we used it for the proof of abstract concavity of topical functions defined on semimodules (see [2], Theorem 3.2).

For an easy reference we present the following definition of the extension of K from [26].

Let $K = (K, \oplus, \otimes)$ be a b-complete idempotent semifield which has no greatest element. Recall that (see [26]) we adjoin to K an outside element, which we denote by $\top$, and extend the canonical order $\leq$ and the addition $\oplus$ from K to an (canonical) order $\leq$ and an addition $\oplus$ on $\overline{K} := K \cup \{\top\}$ by

$$\varepsilon \leq \alpha \leq \top, \quad \forall\, \alpha \in \overline{K},$$

and

$$\alpha \oplus \top = \top \oplus \alpha = \top, \quad \forall\, \alpha \in \overline{K}.$$

Hence the equivalence $\alpha \leq \beta \Longleftrightarrow \alpha \oplus \beta = \beta$ remains valid for all $\alpha, \beta \in \overline{K}$. Furthermore, we extend the multiplication $\otimes$ from K to $\overline{K} := K \cup \{\top\}$ to two multiplications $\dot{\otimes}$ and $\otimes$ by the following rules:

$$\alpha \dot{\otimes} \beta = \alpha \otimes \beta, \quad \forall\ \alpha, \beta \in K,$$
$$\alpha \dot{\otimes} \top = \top \dot{\otimes} \alpha = \top, \quad \forall\ \alpha \in \overline{K},$$
$$\alpha \otimes \top = \top \otimes \alpha = \top, \quad \forall\ \alpha \in \overline{K} \setminus \{\varepsilon\},$$
$$\alpha \otimes \varepsilon = \varepsilon \otimes \alpha = \varepsilon, \quad \forall\ \alpha \in \overline{K}.$$

We often denote the extended product $\otimes$ also by concatenation, which will cause no confusion.

Now, suppose that the Assumption (A0) holds. Consider the function $\ell : X \times X \backslash \{\perp\} \longrightarrow K$ defined by:

$$\ell(x, y) = \inf\{\lambda \in K : x \leq \lambda y\}, \quad \forall\ x \in X,\ \forall\ y \in X \setminus \{\perp\}.$$

Equivalently, one can define ℓ by

$$x \leq \lambda y \Longleftrightarrow \ell(x, y) \leq \lambda, \quad \forall\ x \in X,\ \forall\ y \in X \setminus \{\perp\} \tag{2.1}$$

(with the convention $\lambda \perp = \lambda \otimes \perp = \perp,\ \forall\ \lambda \in K$).

Note that since the Assumption (A0) holds, then $\ell(x, y) \in K$ $(x \in X,\ y \in X \setminus \{\perp\})$, and so $\inf = \min$.

For each $y \in X \setminus \{\perp\}$, define the function $\ell_y : X \longrightarrow K$ by $\ell_y(x) = \ell(x, y)$ for all $x \in X$. For each $x, y \in X$ with $y \neq \perp$ and each $\lambda \in K$, the function ℓ_y has the following properties (see [2]).

$$\ell_y(\lambda x) = \lambda \ell_y(x). \tag{2.2}$$

$$\ell_y(y) = e. \tag{2.3}$$

$$x \leq \ell_y(x) y. \tag{2.4}$$

$$\ell_{\alpha y}(x) = \alpha^{-1} \ell_y(x), \quad \forall\ \alpha \in K \backslash \{\varepsilon\}. \tag{2.5}$$

$$\text{If} x_1, x_2 \in X \textit{ with } x_1 \leq x_2 \text{ and } y \in X \setminus \{\perp\}, \text{ then, } \ell_y(x_1) \leq \ell_y(x_2). \tag{2.6}$$

$$\text{If } y_1, y_2 \in X \setminus \{\perp\} \text{ with } y_1 \leq y_2, \text{ then, } \ell_{y_1}(x) \geq \ell_{y_2}(x), \quad \forall\ x \in X. \tag{2.7}$$

It follows from (2.2) and (2.6) that, for every $y \in X \backslash \{\perp\}$, the function ℓ_y is topical. Let

$$L := \{\ell_y : y \in X \backslash \{\perp\}\}. \tag{2.8}$$

We call L the set of elementary topical functions.

Remark 2.1. If we define the function $\psi : X \setminus \{\perp\} \longrightarrow L$ by $\psi(y) := \ell_y$ for each $y \in X \setminus \{\perp\}$, then, ψ is bijective (one-to-one and onto) and $\psi(\lambda y) = \lambda^{-1}\psi(y)$ for all $y \in X \setminus \{\perp\}$ and all $\lambda \in K \setminus \{\varepsilon\}$.

Example 2.1. Let $X := \mathbb{R}^k_{\max} = \{\mathbb{R} \cup (-\infty)\}^k$, $K := \mathbb{R}_{\max} = \mathbb{R} \cup \{-\infty\}$, $\oplus := \max$, $\otimes := +$, where $\mathbb{R}$ is the set of all real numbers. Put, $\overline{-\infty} := (-\infty, ..., -\infty)$ (k *times*). So, we have $\varepsilon = -\infty$, $e = 0$, $x \oplus y = \max\{x, y\}$. Thus, one has $x \leq y \Longleftrightarrow x_i \leq y_i$, $\forall\ 1 \leq i \leq k$, $\forall\ x \in \mathbb{R}^k_{\max}$, $\forall\ y \in \mathbb{R}^k_{\max} \backslash \{\inf \mathbb{R}^k_{\max}\} = \mathbb{R}^k_{\max} \backslash \{\overline{-\infty}\}$. Therefore, we have

$$\begin{aligned}
\ell_y(x) &= \min\{\lambda \in \mathbb{R}_{\max} : (x_1, ..., x_k) \leq \lambda \otimes (y_1, ..., y_k)\} \\
&= \min\{\lambda \in \mathbb{R}_{\max} : (x_1, ..., x_k) \leq \lambda \mathbb{1}_k + (y_1, ..., y_k)\} \\
&= \min\{\lambda \in \mathbb{R}_{\max} : x_i \leq \lambda + y_i,\ \forall\ 1 \leq i \leq k\ \} \\
&= \min\{\lambda \in \mathbb{R}_{\max} : x_i \otimes (-y_i) \leq \lambda,\ \forall\ 1 \leq i \leq k\} \\
&= \max_{1 \leq i \leq k} \{x_i \otimes (-y_i)\},
\end{aligned}$$

where $\mathbb{1}_k := (1, 1, \ldots, 1) \in \mathbb{R}^k$.

So, if $y \in \mathbb{R}^k_{\max} \backslash \{\overline{-\infty}\}$, then one has

$$\ell_y(x) = \begin{cases} \max_{1 \leq i \leq k}\{x_i \otimes (-y_i)\}, & \text{if } x \in \mathbb{R}^k, \\ -\infty, & \text{if } x \in \mathbb{R}^k_{\max} \backslash \mathbb{R}^k. \end{cases}$$

In the sequel, for an easy reference we give the following properties of $\overline{K}$ from [26].

For the inverses in $\overline{K}$ with respect to $\otimes$ we make the following conventions

$$\varepsilon^{-1} := \top, \quad \top^{-1} := \varepsilon.$$

Whence, by the above one has

$$\varepsilon^{-1}\varepsilon = \top\varepsilon = \varepsilon \neq e, \quad \varepsilon^{-1}\dot{\otimes}\varepsilon = \top\dot{\otimes}\varepsilon = \top \neq e,$$
$$\top^{-1}\top = \varepsilon\top = \varepsilon \neq e, \quad \top^{-1}\dot{\otimes}\top = \varepsilon\dot{\otimes}\top = \top \neq e.$$

We call the set $\overline{K} := K \cup \{\top\}$ endowed with the operations $\oplus$, $\otimes$ and $\dot{\otimes}$ the minimal enlargement of K.

Recall that (see [26, Remark 3]) the product $\dot{\otimes}$ on $\overline{K}$ is associative. Also, one has

$$\alpha \otimes e = e \otimes \alpha = \alpha, \quad \forall\ \alpha \in \overline{K},$$

and

$$\alpha \dot{\otimes} e = e \dot{\otimes} \alpha = \alpha, \quad \forall\ \alpha \in \overline{K}.$$

That is, e is the unit element of $\overline{K}$ for both products $\otimes$ and $\dot{\otimes}$. By the definition of a semimodule X over K we have,

$$\lambda\perp = \lambda\dot{\otimes}\perp := \perp, \quad \forall\, \lambda \in K.$$

Now, we extend the above formula to $\lambda = \top$ by defining (see [26, Definition 4])

$$\top\perp = \top\otimes\perp := \perp,$$

and define

$$\lambda\dot{\otimes}\perp := \lambda \otimes \perp = \perp, \quad \forall\, \lambda \in \overline{K}.$$

In the sequel, we give a definition for an extended valued homogeneous function $f : X \longrightarrow \overline{K}$ (see [26]).

An extended valued function $f : X \longrightarrow \overline{K}$ is called homogeneous if

$$f(\lambda x) = \lambda f(x), \quad \forall\, x \in X,\ \forall\, \lambda \in K.$$

An extended valued function $f : X \longrightarrow \overline{K}$ is called topical if f is homogeneous and f is increasing (i.e., if $x, y \in X$ and $x \le y \Longrightarrow f(x) \le f(y)$).

We also define [2] an extended valued elementary function $\widetilde{\ell_y} : X \longrightarrow \overline{K}$ by

$$\widetilde{\ell_y}(x) := \inf\{\lambda \in K : x \le \lambda y\}, \quad \forall\, x, y \in X,$$

(with conventions $\inf\emptyset := \top$ and $\inf K = \varepsilon$). It is easy to check that $\widetilde{\ell_y}$ has all properties of ℓ_y (see (2.1)–(2.7)). It is worth noting that under the Assumption (A0) one has $\widetilde{\ell_y}(x) = \ell_y(x)$ for all $x \in X$ and all $y \in X \setminus \{\perp\}$. Moreover, one has

$$\widetilde{\ell_\perp}(x) = \inf\{\lambda \in K : x \le \lambda\perp\} = \begin{cases} \inf K = \varepsilon, & \text{if } x = \perp, \\ \inf\emptyset = \top, & \text{if } x \ne \perp, \end{cases} \tag{2.9}$$

and

$$\widetilde{\ell_y}(\perp) = \inf\{\lambda \in K : \perp \le \lambda y\} = \inf K = \varepsilon, \quad \forall\, y \in X. \tag{2.10}$$

It is easy to see that for each $y \in X$, $\widetilde{\ell_y}$ is a topical function.

Lemma 2.1. *([2], Lemma 3.3) Under the Assumption (A0) we have*

$$\widetilde{\ell_{\alpha y}}(x) = \begin{cases} \alpha^{-1}\dot{\otimes}\widetilde{\ell_y}(x), & \textit{if } x \in X \setminus \{\perp\},\ y \in X \textit{ and } \alpha \in K, \\ \alpha^{-1}\dot{\otimes}\widetilde{\ell_y}(x), & \textit{if } x = \perp,\ y \in X \textit{ and } \alpha \in K \setminus \{\varepsilon\}, \\ \alpha^{-1}\otimes\widetilde{\ell_y}(x), & \textit{if } x = \perp,\ y \in X \textit{ and } \alpha = \varepsilon. \end{cases}$$

Now, let

$$\widetilde{L} := \{\widetilde{\ell}_y : y \in X\}. \tag{2.11}$$

We call $\widetilde{L}$ the set of extended valued elementary topical functions.

Remark 2.2. In view of Remark 2.1, There is a one-to-one correspondence between X and $\widetilde{L}$ (see also, [2]).

Theorem 2.1. *(See [2], Theorem 3.3) Let $f : X \longrightarrow \overline{K}$ be a function. Suppose that the Assumption (A0) holds. Then the following assertions are equivalent:*

(i) *f is topical.*
(ii) *$f(\perp) = \varepsilon$ and $f(x) \leq \lambda f(y)$ for all $x, y \in X$ and all $\lambda \in K$ such that $x \leq \lambda y$.*
(iii) *$f(\perp) = \varepsilon$ and $f(x) \leq \widetilde{\ell}_y(x) \dot{\otimes} f(y)$ for all $x, y \in X$.*

Theorem 2.2. *(See [2], Theorem 3.4) Let $f : X \longrightarrow \overline{K}$ be a function and $\widetilde{L}$ be the set defined by (2.11). Suppose that the Assumption (A0) holds. Then the following assertions are equivalent:*

(a) *f is topical.*
(b) *There exists a set $\widetilde{L_0} \subseteq \widetilde{L}$ such that*

$$f(x) = \inf_{\widetilde{\ell}_y \in \widetilde{L_0}} \widetilde{\ell}_y(x), \quad (x \in X).$$

In this case, one can take $\widetilde{L_0} := \{\widetilde{\ell}_y \in \widetilde{L} : f(y) \leq e\}$. So, f is topical if and only if f is $\widetilde{L}$-concave.

Recall (see [20]) that for a function $f : X \longrightarrow \overline{K}$, define the upper support set of f with respect to $\widetilde{L}$ by

$$\text{supp}_u(f, \widetilde{L}) := \{y \in X : \widetilde{\ell}_y(x) \geq f(x), \ \forall\, x \in X\},$$

(see also, Remark 2.2).

Proposition 2.1. *(See [2], Lemma 3.4) Let $f : X \longrightarrow \overline{K}$ be a topical function. Suppose that the Assumption (A0) holds. Then,*

$$\text{supp}_u(f, \widetilde{L}) := \{y \in X : f(y) \leq e\}.$$

3. Characterizations of the superdifferential of topical functions on semimodules

In this section, we first define the $\tilde{L}$-superdifferential for extended valued topical functions defined on a semimodule, and then, we give characterizations of the $\tilde{L}$-superdifferential for this class of functions.

Definition 3.1. For a function $f : X \longrightarrow \overline{K}$, define the $\tilde{L}$-superdifferential of f at a point $x_0 \in X \setminus \{\perp\}$, by

$$\partial^+_{\tilde{L}} f(x_0) := \{y \in X : \widetilde{\ell_y}(x) \dot{\otimes} [\widetilde{\ell_y}(x_0)]^{-1} \dot{\otimes} f(x_0) \geq f(x),\ \forall\ x \in X\}.$$

The following result gives a characterization of the $\tilde{L}$-superdifferential of an extended valued topical function.

Theorem 3.1. *Let $f : X \longrightarrow \overline{K}$ be a topical function, and let $x_0 \in X \setminus \{\perp\}$ be such that $f(x_0) \in K$. Assume that the Assumption (A0) holds. Then,*

$$\partial^+_{\tilde{L}} f(x_0) = \{y \in X : f(y) = [\widetilde{\ell_y}(x_0)]^{-1} \dot{\otimes} f(x_0)\}.$$

Proof. Put

$$D := \{y \in X : f(y) = [\widetilde{\ell_y}(x_0)]^{-1} \dot{\otimes} f(x_0)\}.$$

Let $y \in D$ be arbitrary. Then,

$$f(y) = [\widetilde{\ell_y}(x_0)]^{-1} \dot{\otimes} f(x_0). \tag{3.1}$$

Now, since f is a topical function, it follows from Theorem 2.1 (the implication (i) $\Longrightarrow$ (iii)) that

$$f(\perp) = \varepsilon \text{ and } f(x) \leq \widetilde{\ell_y}(x) \dot{\otimes} f(y),\ \forall\ x \in X. \tag{3.2}$$

If $y = \perp$, since $x_0 \neq \perp$, then, in view of (2.9), $\widetilde{\ell_y}(x_0) = \top$. Thus, by (3.2), (2.9) and (2.10) one has

$$\widetilde{\ell_y}(x) \dot{\otimes} [\widetilde{\ell_y}(x_0)]^{-1} \dot{\otimes} f(x_0) = \begin{cases} \varepsilon = f(x), & x = \perp, \\ \top \geq f(x), & x \neq \perp,\ \forall\ x \in X. \end{cases}$$

So, $y \in \partial^+_{\tilde{L}} f(x_0)$. If $y \neq \perp$, it follows from the definition of $\widetilde{\ell_y}$ and the fact that the Assumption (A0) holds, $\widetilde{\ell_y} = \ell_y$. Hence, since $x_0 \neq \perp$, we conclude that $\widetilde{\ell_y}(x_0) \neq \varepsilon$, because if $\widetilde{\ell_y}(x_0) = \varepsilon$, then, $\ell_y(x_0) = \varepsilon$. Thus, in view of (2.1), one has $x_0 \leq \varepsilon y = \perp$. That is, $x_0 = \perp$, which is a contradiction (note that $\varepsilon x = \perp$ for all $x \in X$). So, $\widetilde{\ell_y}(x_0)$ is invertible in K. Therefore, in view of (3.1) and (3.2),

$$\widetilde{\ell_y}(x)\dot{\otimes}[\widetilde{\ell_y}(x_0)]^{-1}\dot{\otimes}f(x_0) = \widetilde{\ell_y}(x)\dot{\otimes}f(y) \geq f(x), \ \forall\ x \in X.$$

This implies that $y \in \partial^+_{\tilde{L}} f(x_0)$. Thus, $D \subseteq \partial^+_{\tilde{L}} f(x_0)$.

Conversely, let $y \in \partial^+_{\tilde{L}} f(x_0)$ be arbitrary. Then, by Definition 3.1, one has

$$\widetilde{\ell_y}(x)\dot{\otimes}[\widetilde{\ell_y}(x_0)]^{-1}\dot{\otimes}f(x_0) \geq f(x), \ \forall\ x \in X. \tag{3.3}$$

If $y = \perp$, since $x_0 \neq \perp$, it follows from (2.9) that $\widetilde{\ell_y}(x_0) = \top$. Since f is a topical function, by (3.2), we get $f(y) = \varepsilon$. Hence,

$$[\widetilde{\ell_y}(x_0)]^{-1}\dot{\otimes}f(x_0) = \varepsilon = f(y).$$

That is, $y \in D$. Now, assume that $y \neq \perp$. Thus, by the definition of $\widetilde{\ell_y}$ and the fact that the Assumption (A0) holds, $\widetilde{\ell_y} = \ell_y$. This together with $x_0 \neq \perp$ implies that $\widetilde{\ell_y}(x_0) \neq \varepsilon$. Therefore, $\widetilde{\ell_y}(x_0)$ is invertible in K. So, in view of (3.3) with $x := y \neq \perp$, and by using (3.2) with $x := x_0$, one has

$$[\widetilde{\ell_y}(x_0)]^{-1}\dot{\otimes}f(x_0) \geq f(y) \geq [\widetilde{\ell_y}(x_0)]^{-1}\dot{\otimes}f(x_0),$$

and so, $[\widetilde{\ell_y}(x_0)]^{-1}\dot{\otimes}f(x_0) = f(y)$. That is, $y \in D$, which completes the proof. □

Remark 3.1. It is worth noting that under the hypotheses of Theorem 3.1, if $y \in \partial^+_{\tilde{L}} f(x_0)$, then, $f(y) \neq \top$ and $\widetilde{\ell_y}(x_0) \neq \varepsilon$, because $f(x_0) \in K$ and $x_0 \in X \setminus \{\perp\}$.

In the following, we give some properties of the $\tilde{L}$-superdifferential of a topical function. First, recall that by the definition of a semimodule X over K, one has

$$\lambda\perp = \lambda\dot{\otimes}\perp := \perp, \quad \forall\ \lambda \in K. \tag{3.4}$$

Proposition 3.1. *Let $f : X \longrightarrow \overline{K}$ be a topical function, and let $x_0 \in X \setminus \{\perp\}$ be such that $f(x_0) \in K$. Assume that the Assumption (A0) holds. If $y \in \partial^+_{\tilde{L}} f(x_0)$, then, $\lambda y \in \partial^+_{\tilde{L}} f(x_0)$ for all $\lambda \in K$ (for $y = \perp$, see (3.4)).*

Proof. Suppose that $\lambda \in K$ and $y \in \partial^+_{\tilde{L}} f(x_0)$ (note that by Remark 3.1, one has $f(y) \neq \top$ and $\widetilde{\ell_y}(x_0) \neq \varepsilon$). Then, in view of Theorem 3.1,

$$f(y) = [\widetilde{\ell_y}(x_0)]^{-1}\dot{\otimes}f(x_0).$$

This together with Lemma 2.1 and the fact that f is homogeneous implies that

$$\begin{aligned} f(\lambda y) &= \lambda f(y) \\ &= \lambda\otimes f(y) \end{aligned}$$

$$
\begin{aligned}
&= \lambda \dot{\otimes} [\widetilde{\ell_y}(x_0)]^{-1} \dot{\otimes} f(x_0) \\
&= [\lambda^{-1} \dot{\otimes} \widetilde{\ell_y}(x_0)]^{-1} \dot{\otimes} f(x_0) \\
&= [\widetilde{\ell_{\lambda y}}(x_0)]^{-1} \dot{\otimes} f(x_0).
\end{aligned}
$$

Again, in view of Theorem 3.1, $\lambda y \in \partial^+_{\tilde{L}} f(x_0)$. □

Definition 3.2. Let A be a subset of X and $\lambda \in K$ be arbitrary. Define

$$\lambda A = \lambda \otimes A := \{\lambda x = \lambda \otimes x : x \in A\}. \tag{3.5}$$

If $\perp \in A$, then, for the definition, see (3.4).

Proposition 3.2. *Let $f : X \longrightarrow \overline{K}$ be a topical function, and let $x_0 \in X \setminus \{\perp\}$ be such that $f(x_0) \in K$. Assume that the Assumption (A0) holds. Then,*

$$\partial^+_{\tilde{L}} f(x_0) = \partial^+_{\tilde{L}} f(\lambda x_0) = \lambda \partial^+_{\tilde{L}} f(x_0), \ \forall \ \lambda \in K \backslash \{\varepsilon\}.$$

(See Definition 3.2.)

Proof. Let $\lambda \in K \setminus \{\varepsilon\}$ be arbitrary. Suppose that $y \in \partial^+_{\tilde{L}} f(x_0)$. Consider two possible cases:

Case (i). If $y = \perp$, then, in view of (3.4) and Definition 3.2, one has

$$y = \lambda \perp = \lambda y \in \lambda \partial^+_{\tilde{L}} f(x_0).$$

Case (ii). Assume that $y \neq \perp$. Thus, by Proposition 3.1, $\lambda^{-1} y \in \partial^+_{\tilde{L}} f(x_0)$. So, $y \in \lambda \partial^+_{\tilde{L}} f(x_0)$.

Therefore, in any case, we have $y \in \lambda \partial^+_{\tilde{L}} f(x_0)$. Now, assume that $y' \in \lambda \partial^+_{\tilde{L}} f(x_0)$. Then by Definition 3.2 there exists $y \in \partial^+_{\tilde{L}} f(x_0)$ such that $y' = \lambda y$. In view of Proposition 3.1 we conclude that $y' \in \partial^+_{\tilde{L}} f(x_0)$. That is,

$$\partial^+_{\tilde{L}} f(x_0) = \lambda \partial^+_{\tilde{L}} f(x_0), \ \forall \ \lambda \in K \backslash \{\varepsilon\}.$$

Furthermore, assume that $y \in \partial^+_{\tilde{L}} f(\lambda x_0)$. Then, in view of Theorem 3.1,

$$[\widetilde{\ell_y}(\lambda x_0)]^{-1} \dot{\otimes} f(\lambda x_0) = f(y).$$

Since f and $\widetilde{\ell_y}$ are topical functions, so by using the properties $\dot{\otimes}$, it follows that

$$[\widetilde{\ell_y}(x_0)]^{-1} \dot{\otimes} f(x_0) = f(y).$$

This together with Theorem 3.1 implies that $y \in \partial^+_{\tilde{L}} f(x_0)$. Similarly, one can show that $\partial^+_{\tilde{L}} f(x_0) \subseteq \partial^+_{\tilde{L}} f(\lambda x_0)$, which completes the proof. □

Lemma 3.1. *Let $f, g : X \longrightarrow \overline{K}$ be topical functions, and let $x_0 \in X \backslash \{\perp\}$ be such that $f(x_0), g(x_0) \in K \backslash \{\varepsilon\}$. Assume that the Assumption (A0) holds. Let $y_0 := [f(x_0)]^{-1}[g(x_0)]^{-1}x_0$. Then, $y_0 \in \partial^+_{\tilde{L}} f(x_0) \cap \partial^+_{\tilde{L}} g(x_0)$.*

Proof. First, note that $y_0 \neq \perp$, because $x_0 \in X \setminus \{\perp\}$. Thus, $\widetilde{\ell_{y_0}} = \ell_{y_0}$. Therefore, by using properties of ℓ_{y_0} and the fact that f is a topical function, one has

$$\begin{aligned} f(y_0) &= f([f(x_0)]^{-1}[g(x_0)]^{-1}x_0) \\ &= [f(x_0)]^{-1}[g(x_0)]^{-1}f(x_0) \\ &= [f(x_0)]^{-1} \otimes [g(x_0)]^{-1} \otimes e \otimes f(x_0) \\ &= \left([f(x_0)]^{-1}\dot{\otimes}[g(x_0)]^{-1}\dot{\otimes}[\ell_{x_0}(x_0)]^{-1}\right)\dot{\otimes}f(x_0) \\ &= [\widetilde{\ell_{y_0}}(x_0)]^{-1}\dot{\otimes}f(x_0). \end{aligned} \tag{3.6}$$

Similarly,

$$g(y_0) = [\widetilde{\ell_{y_0}}(x_0)]^{-1}\dot{\otimes}g(x_0). \tag{3.7}$$

Hence, in view of Theorem 3.1, it follows from (3.6) and (3.7) that $y_0 \in \partial^+_{\tilde{L}} f(x_0) \cap \partial^+_{\tilde{L}} g(x_0)$. $\square$

Example 3.1. Let $X := \mathbb{R}_{\max} := \mathbb{R} \cup \{-\infty\}$, $K := R_{\max}$, $\otimes := +$ and $\oplus := \max$, where $\mathbb{R}$ is the set of real numbers. Assume that $f : X \longrightarrow K$ is defined by $f(x) = x$ for all $x \in X$. It is clear that f is a topical function, and also $\varepsilon := -\infty$, $e := 0$ and $\perp = -\infty$. Now, let $x_0 \in X\backslash\{\perp\}$ be such that $f(x_0) \in K$. Then,

$$\partial^+_{\tilde{L}} f(x_0) = X.$$

Note that $\perp = -\infty = \varepsilon$, and $\dot{\otimes} = \otimes$. It is easy to see that the Assumption (A0) holds. Therefore, in view of Theorem 3.1, we have

$$\begin{aligned} \partial^+_{\tilde{L}} f(x_0) &= \{y \in X : f(y) = [\widetilde{\ell_y}(x_0)]^{-1}\dot{\otimes}f(x_0)\} \\ &= \{y \in X : y = -\widetilde{\ell_y}(x_0) + x_0\}. \end{aligned} \tag{3.8}$$

It is not difficult to show that

$$\widetilde{\ell_y}(x_0) = \begin{cases} x_0 - y, & y \neq \perp, \\ +\infty, & y = \perp, \ \forall\, y \in X. \end{cases} \tag{3.9}$$

Hence, it follows from (3.8) and (3.9) that $\partial^+_{\tilde{L}} f(x_0) = X$.

4. Characterizations of minimal elements of topical functions

In this section, we first characterize minimal elements of the upper support set of topical functions defined on a semimodule with values in a semifield. Finally, as an application, we give a necessary and sufficient condition for global maximum of the difference of two strictly topical functions.

In [1,3,26,23,24], it has been studied the topologies on lattice ordered groups, b-complete semimodules and b-complete semifields. Let X be a b-complete idempotent semimodule over a b-complete idempotent semifield K, where X and K are equipped with certain topologies which were given in [1,3,26,23,24].

Now, we make the following assumptions (also, see [2,26,24]).

Assumption (C1). We assume that the idempotent addition $\oplus : X \times X \longrightarrow X$ is continuous.

Assumption (C2). For each $x \in X$, the function $u_x : \lambda \in K \longrightarrow \lambda x \in X$ is continuous. That is, for any $x \in X$, $\lambda \in K$ and any net $\{\lambda_k\} \subset K$ such that $\lambda_k \longrightarrow \lambda$, we have $\lambda_k x \longrightarrow \lambda x$.

Assumption (C1). Let D be an arbitrary subset of K and $\delta \in K$ with $\delta := \inf D$, then there exists a net $\{d_k\} \subset D$ such that $d_k \longrightarrow \delta$.

Definition 4.1. Let $U \subseteq \tilde{L}$ be a set of functions. A function $f \in U$ is called a minimal element of the set U, if $\tilde{f} \in U$ is such that $\tilde{f}(x) \leq f(x)$ for all $x \in X$, then, $\tilde{f}(x) = f(x)$ for all $x \in X$.

Definition 4.2. A function $f : X \longrightarrow \overline{K}$ is called strictly topical if f is homogeneous and strictly increasing (the later means that if $x < y \Longrightarrow f(x) < f(y),\ x, y \in X$).

Lemma 4.1. *Let $f : X \longrightarrow \overline{K}$ be a topical function, and let the Assumption (A0) hold. Let $y \in X \setminus \{\perp\}$. If $\widetilde{\ell_y} \in \tilde{L}$ is a minimal element of $\mathrm{supp}_u(f, \tilde{L})$, then, $f(y) \neq \top,\ \varepsilon$.*

Proof. Assume that $f(y) = \top$. Since $\top \nleq e$, in view of Proposition 2.1 one has $\widetilde{\ell_y} \notin \mathrm{supp}_u(f, \tilde{L})$. This is a contradiction. Now, suppose that $f(y) = \varepsilon$. Let $\lambda \in K \setminus \{\varepsilon\}$ be such that $\lambda < e$. Since f is topical, thus,

$$f(\lambda^{-1} y) = \lambda^{-1} f(y) = \lambda^{-1} \otimes \varepsilon = \varepsilon \leq e.$$

This together with Proposition 2.1 implies that $\widetilde{\ell_{\lambda^{-1}y}} \in \mathrm{supp}_u(f, \tilde{L})$. So, it follows from Lemma 2.1 that

$$\widetilde{\ell_{\lambda^{-1}y}}(x) = \lambda \dot{\otimes} \widetilde{\ell_y}(x) \leq e \dot{\otimes} \widetilde{\ell_y}(x) = \widetilde{\ell_y}(x),\ \forall\ x \in X.$$

Since $\widetilde{\ell_y}$ is a minimal element of $\operatorname{supp}_u(f, \tilde{L})$, we conclude that

$$\widetilde{\ell_y}(x) = \widetilde{\ell_{\lambda^{-1}y}}(x) = \lambda \dot{\otimes} \widetilde{\ell_y}(x), \ \forall\, x \in X. \tag{4.1}$$

Put $x := y$ in (4.1). Therefore, $\lambda = e$, which contradicts $\lambda < e$. It is worth noting that, since $y \in X \setminus \{\perp\}$, we have $\widetilde{\ell_y}(y) = \ell_y(y) = e$. $\square$

Proposition 4.1. *Let $f : X \longrightarrow \overline{K}$ be a topical function, and let the Assumption (A0) hold. Let $y \in X \setminus \{\perp\}$. If $\widetilde{\ell_y} \in \tilde{L}$ is a minimal element of* $\operatorname{supp}_u(f, \tilde{L})$*, then, $f(y) = e$.*

Proof. Since $\widetilde{\ell_y} \in \operatorname{supp}_u(f, \tilde{L})$, it follows from Proposition 2.1 that $f(y) \le e$. But, by Lemma 4.1, $f(y) \ne \top,\ \varepsilon$. Then, $f(y)$ is invertible in K. Now, set $y' := [f(y)]^{-1}y$, and so, $f(y') = e$. Thus, in view of Proposition 2.1 one has $\widetilde{\ell_{y'}} \in \operatorname{supp}_u(f, \tilde{L})$. Since $f(y) \le e$, by using Lemma 2.1 we conclude that

$$\widetilde{\ell_{y'}}(x) = f(y) \dot{\otimes} \widetilde{\ell_y}(x) \le e \dot{\otimes} \widetilde{\ell_y}(x) = \widetilde{\ell_y}(x), \ \forall\, x \in X. \tag{4.2}$$

Since, by the hypothesis $\widetilde{\ell_y}$ is a minimal element of $\operatorname{supp}_u(f, \tilde{L})$, it follows from (4.2) that

$$\widetilde{\ell_{y'}}(x) = \widetilde{\ell_y}(x), \ \forall\, x \in X.$$

This together with (4.2) implies that

$$f(y) \dot{\otimes} \widetilde{\ell_y}(x) = \widetilde{\ell_y}(x), \ \forall\, x \in X. \tag{4.3}$$

Put $x := y$ in (4.3). So, it follows that $f(y) = e$, which completes the proof. $\square$

Lemma 4.2. *Let $f : X \longrightarrow \overline{K}$ be a topical function, and let the Assumptions (A0), (C1), (C2) and (C1) hold. Let $y \in X$ be such that $f(y) = e$. Assume that there exists $\widetilde{\ell_{y'}} \in \operatorname{supp}_u(f, \tilde{L})$ such that $\widetilde{\ell_{y'}}(x) \le \widetilde{\ell_y}(x)$ for all $x \in X$. Then, $y \le y'$ and $f(y') = e$.*

Proof. In view of Theorem 2.1, since f is a topical function and $f(y) = e$, it follows that $y \in X \setminus \{\perp\}$. But, by the hypothesis one has $\widetilde{\ell_{y'}} \in \operatorname{supp}_u(f, \tilde{L})$, then,

$$\widetilde{\ell_{y'}}(x) \ge f(x), \ \forall\, x \in X, \tag{4.4}$$

and also, since f is topical, by Proposition 2.1, we have $f(y') \le e$. Therefore, by the hypothesis and (4.4),

$$e = f(y) \le \widetilde{\ell_{y'}}(y) \le \widetilde{\ell_y}(y) = e.$$

This implies that $\widetilde{\ell_{y'}}(y) = e$. Thus, by the definition of $\widetilde{\ell_y}$ and the Assumption (C1), there exists a net $\{\lambda_n\} \subset K$ such that $\lambda_n \longrightarrow e$ and $y \le \lambda_n y'$ for all n. By the

Assumptions (C1) and (C2), we conclude that $y \leq ey' = y'$. Since f is increasing, so, $f(y) \leq f(y')$. This together with $f(y) = e$ and the fact that $f(y') \leq e$ implies that $f(y') = e$. □

Remark 4.1. The converse statement to Proposition 4.1 is not valid. Indeed, let $X := \mathbb{R}_{\max} := \mathbb{R} \cup \{-\infty\}$, $K := R_{\max}$, $\otimes := +$ and $\oplus := \max$, where $\mathbb{R}$ is the set of real numbers. Assume that $f : X \longrightarrow K$ is defined by $f(x) = x$ for all $x \in X$. It is clear that f is a topical function, and also $\varepsilon := -\infty$, $e := 0$ and $\perp = -\infty$. It is easy to see that the Assumptions (A0) holds. In view of Proposition 2.1, one has

$$\begin{aligned} \operatorname{supp}_u(f, \tilde{L}) &= \{y \in X : f(y) \leq e\} \\ &= \{y \in X : y \leq 0\} \\ &= [-\infty, 0]. \end{aligned}$$

It is clear that $\widetilde{\ell_0} \in \operatorname{supp}_u(f, \tilde{L})$ and $f(0) = 0 = e$. But, $\operatorname{supp}_u(f, \tilde{L}) \setminus \{\perp\} = (-\infty, 0]$. Therefore, the minimal element of $supp_u(f, \tilde{L}) \setminus \{\perp\}$ does not exist.

Now, we show that under extra conditions the converse statement to Proposition 4.1 holds. In fact, in this case, we also assume that f is a strictly topical function.

Theorem 4.1. *Suppose that the Assumptions (A0), (C1), (C2) and (C1) hold. Let $f : X \longrightarrow \overline{K}$ be a strictly topical function, and let $\widetilde{\ell_y} \in \operatorname{supp}_u(f, \tilde{L})$ be such that $y \in X \setminus \{\perp\}$. Then, $\widetilde{\ell_y}$ is a minimal element of $\operatorname{supp}_u(f, \tilde{L})$ if and only if $f(y) = e$.*

Proof. Due to Proposition 4.1, we only prove that if $f(y) = e$, then, $\widetilde{\ell_y}$ is a minimal element of $\operatorname{supp}_u(f, \tilde{L})$. To end this, let $\widetilde{\ell_{y'}} \in \operatorname{supp}_u(f, \tilde{L})$ be such that $\widetilde{\ell_{y'}}(x) \leq \widetilde{\ell_y}(x)$ for all $x \in X$. Then, in view of Lemma 4.2, we have $y \leq y'$ and $f(y') = e$. Therefore, $f(y) = f(y')$. Since f is strictly increasing, we conclude from Definition 4.2 that $y = y'$, and hence, $\widetilde{\ell_{y'}}(x) = \widetilde{\ell_y}(x)$ for all $x \in X$, which completes the proof. □

Proposition 4.2. *Suppose that the Assumptions (A0), (C1), (C2) and (C1) hold. Let $f : X \longrightarrow \overline{K}$ be a strictly topical function. Then, for each $\widetilde{\ell_y} \in \operatorname{supp}_u(f, \tilde{L})$ with $f(y) \neq \varepsilon$, there exists a minimal element $\widetilde{\ell_{y_0}} \in \operatorname{supp}_u(f, \tilde{L})$ such that $\widetilde{\ell_{y_0}}(x) \leq \widetilde{\ell_y}(x)$ for all $x \in X$. In this case, one can take $y_0 := [f(y)]^{-1}y$ (it is worth noting that, since $f(y) \leq e$, $f(y) \neq \top$).*

Proof. Let $\widetilde{\ell_y} \in \operatorname{supp}_u(f, \tilde{L})$ with $f(y) \neq \varepsilon$. Thus, by Proposition 2.1, we have $f(y) \leq e$. Put $y_0 := [f(y)]^{-1}y$. Since f is topical, $f(y_0) = [f(y)]^{-1}f(y) = e$, and so, in view of Proposition 2.1, $\widetilde{\ell_{y_0}} \in \operatorname{supp}_u(f, \tilde{L})$. Also, one has $y_0 \in X \setminus \{\perp\}$ because f is topical and $f(y_0) = e$. On the other hand, since $f(y_0) = e$ and f is strictly topical function, it follows from Theorem 4.1 that $\widetilde{\ell_{y_0}}$ is a minimal element $\operatorname{supp}_u(f, \tilde{L})$. Now, since $f(y) \leq e$, by using Lemma 2.1 we conclude that

$$\widetilde{\ell_{y_0}}(x) = f(y)\dot{\otimes}\widetilde{\ell_y}(x) \leq e\dot{\otimes}\widetilde{\ell_y}(x) = \widetilde{\ell_y}(x), \ \forall \ x \in X.$$

Hence, the proof is complete. □

Theorem 4.2. *Suppose that the Assumptions (A0), (C1), (C2) and (C1) hold. Let $f, g : X \longrightarrow \overline{K}$ be strictly topical functions such that $f(x), g(x) \neq \varepsilon$ for all $x \in X \setminus \{\perp\}$. Then the following assertions are equivalent:*

(1) $\operatorname{supp}_u(f, \tilde{L}) \subseteq \operatorname{supp}_u(g, \tilde{L})$.
(2) *For each minimal element $\widetilde{\ell_{y_1}}$ of $\operatorname{supp}_u(f, \tilde{L})$ with $y_1 \neq \perp$ there exists a minimal element $\widetilde{\ell_{y_2}}$ of $\operatorname{supp}_u(g, \tilde{L})$ such that $\widetilde{\ell_{y_2}}(x) \leq \widetilde{\ell_{y_1}}(x)$ for all $x \in X$.*
(3) *$g(x) \leq f(x)$ for all $x \in X$.*

Proof. (1) $\Longrightarrow$ (2). Suppose that $\operatorname{supp}_u(f, \tilde{L}) \subseteq \operatorname{supp}_u(g, \tilde{L})$. Let $\widetilde{\ell_{y_1}} \in \tilde{L}$ be an arbitrary minimal element of $\operatorname{supp}_u(f, \tilde{L})$ with $y_1 \neq \perp$. Then, $\widetilde{\ell_{y_1}} \in \operatorname{supp}_u(g, \tilde{L})$. This together with Proposition 4.2 and the fact that $g(y_1) \neq \varepsilon$ (by the hypothesis because $y_1 \neq \perp$) implies that there exists a minimal element $\widetilde{\ell_{y_2}}$ of $\operatorname{supp}_u(g, \tilde{L})$ such that $\widetilde{\ell_{y_2}}(x) \leq \widetilde{\ell_{y_1}}(x)$ for all $x \in X$.

(2) $\Longrightarrow$ (1). Let $\widetilde{\ell_y} \in \operatorname{supp}_u(f, \tilde{L})$ be arbitrary. Consider two possible cases:

Case (i). If $y = \perp$, then, by (2.9) and the fact that g is topical, one has

$$\widetilde{\ell_y}(x) = \begin{cases} \varepsilon = g(\perp) = g(x), & x = \perp, \\ \top \geq g(x), & x \neq \perp, \ \forall \ x \in X. \end{cases}$$

This implies that $\widetilde{\ell_y} \in \operatorname{supp}_u(g, \tilde{L})$.

Case (ii). Suppose that $y \neq \perp$. Thus, by the hypothesis, $f(y) \neq \varepsilon$, and so, it follows from Proposition 4.2 that there exists a minimal element $\widetilde{\ell_{y_1}}$ of $\operatorname{supp}_u(f, \tilde{L})$ such that

$$\widetilde{\ell_{y_1}}(x) \leq \widetilde{\ell_y}(x), \ \forall \ x \in X. \tag{4.5}$$

This together with $y \neq \perp$ and $\widetilde{\ell_y}(y) = \ell_y(y) = e$ implies that $y_1 \neq \perp$ because if $y_1 = \perp$, then, in view of (4.5) and (2.9) and the fact that $y \neq \perp$, we get $\top = \widetilde{\ell_{y_1}}(y) \leq \widetilde{\ell_y}(y) = e$, which is a contradiction. Hence, $y_1 \neq \perp$, and also $\widetilde{\ell_{y_1}}$ is a minimal element of $\operatorname{supp}_u(f, \tilde{L})$, thus, by the hypothesis (2), there exists a minimal element $\widetilde{\ell_{y_2}}$ of $\operatorname{supp}_u(g, \tilde{L})$ such that

$$\widetilde{\ell_{y_2}}(x) \leq \widetilde{\ell_{y_1}}(x), \ \forall \ x \in X. \tag{4.6}$$

Therefore, it follows from (4.5), (4.6) and the fact that $\widetilde{\ell_{y_2}} \in \operatorname{supp}_u(g, \tilde{L})$,

$$\widetilde{\ell_y}(x) \geq g(x), \ \forall \ x \in X.$$

Hence, $\widetilde{\ell_y} \in \operatorname{supp}_u(g, \tilde{L})$. So, in any case, (1) holds.

(1) $\Longrightarrow$ (3). Assume that $\operatorname{supp}_u(f,\tilde{L}) \subseteq \operatorname{supp}_u(g,\tilde{L})$. Let $y \in X \setminus \{\perp\}$ be arbitrary. In view of the hypothesis we have $f(y) \neq \varepsilon$. Put $y_1 := [f(y)]^{-1}y$. Then, $f(y_1) = e$, and so by Proposition 2.1, $\widetilde{\ell_{y_1}} \in \operatorname{supp}_u(f,\tilde{L})$. Therefore, by the hypothesis (1), one has $\widetilde{\ell_{y_1}} \in \operatorname{supp}_u(g,\tilde{L})$. It follows from Proposition 2.1 that $g(y_1) \leq e$. This together with $y_1 := [f(y)]^{-1}y$ and the fact that g is topical implies that

$$[f(y)]^{-1}g(y) = g(y_1) \leq e,$$

and hence, $g(y) \leq f(y)$. If $y = \perp$, then, since f and g are topical functions, $g(y) = \varepsilon = f(y)$. Thus, $g \leq f$ on X.

(3) $\Longrightarrow$ (1). Suppose that $g(x) \leq f(x)$ for all $x \in X$. Then, by the definition of the upper support set, we conclude that (1) holds. □

Proposition 4.3. *Suppose that the Assumptions (A0), (C1), (C2) and (C1) hold. Let $f, g : X \longrightarrow \overline{K}$ be topical functions and $\lambda \in K \setminus \{\varepsilon\}$. Define $\tilde{f}(x) := \lambda \dot{\otimes} f(x)$ for all $x \in X$. Then, $g(x) \leq \tilde{f}(x)$ for all $x \in X$ if and only if*

$$\operatorname{supp}_u(\tilde{f},\tilde{L}) \subseteq \operatorname{supp}_u(g,\tilde{L}).$$

Proof. It is not difficult to show that if $g(x) \leq \tilde{f}(x)$ for all $x \in X$, then, $\operatorname{supp}_u(\tilde{f},\tilde{L}) \subseteq \operatorname{supp}_u(g,\tilde{L})$. Conversely, let $x \in X$ be arbitrary. If $f(x) = \top$, then, $\tilde{f}(x) = \top \geq g(x)$. Now, assume that $f(x) \neq \top$. Therefore, in view of (2.2), Theorem 2.2, Lemma 2.1 and the fact that f is a topical function, we have

$$\begin{aligned}
\tilde{f}(x) &= \lambda \dot{\otimes} f(x) \\
&= \lambda \otimes f(x) \\
&= f(\lambda \otimes x) \\
&= f(\lambda x) \\
&= \inf_{\widetilde{\ell_y} \in \operatorname{supp}_u(f,\tilde{L})} \widetilde{\ell_y}(\lambda x) \\
&= \inf_{\widetilde{\ell_y} \in \operatorname{supp}_u(f,\tilde{L})} [\lambda \dot{\otimes} \widetilde{\ell_y}(x)] \\
&= \inf_{\widetilde{\ell_y} \in \operatorname{supp}_u(f,\tilde{L})} \widetilde{\ell_{\lambda^{-1}y}}(x) \\
&\geq \inf_{\widetilde{\ell_z} \in \operatorname{supp}_u(g,\tilde{L})} \widetilde{\ell_z}(x) \\
&= g(x).
\end{aligned}$$

It is worth nothing that if $\widetilde{\ell_y} \in \operatorname{supp}_u(f,\tilde{L})$, then, by Lemma 2.1,

$$\begin{aligned}\widetilde{\ell_{\lambda^{-1}y}}(x) &= \lambda\dot{\otimes}\widetilde{\ell_y}(x)\\ &\geq \lambda\dot{\otimes}f(x)\\ &= \tilde{f}(x),\ \forall x \in X.\end{aligned}$$

Thus, by the definition of the upper support set, one has

$$\widetilde{\ell_{\lambda^{-1}y}} \in \operatorname{supp}_u(\tilde{f}, \tilde{L}) \subseteq \operatorname{supp}_u(g, \tilde{L}),$$

where the later inclusion holds by the hypothesis. □

The following theorem has a crucial role for global maximizing of the difference of two strictly topical functions.

Theorem 4.3. *Suppose that the Assumptions (A0), (C1), (C2) and (C1) hold. Let $f, g : X \longrightarrow \overline{K}$ be strictly topical functions such that $f(x), g(x) \neq \varepsilon$ for all $x \in X \setminus \{\perp\}$. Let $\lambda \in K \setminus \{\varepsilon\}$ be fixed and arbitrary. Define the function $\tilde{f} : X \longrightarrow \overline{K}$ by $\tilde{f}(x) := \lambda\dot{\otimes}f(x)$ for all $x \in X$ (note that $\tilde{f}$ is a strictly topical function and $\tilde{f}(x) \neq \varepsilon$ for all $x \in X\setminus\{\perp\}$). Then, $g(x) \leq \tilde{f}(x)$ for all $x \in X$ if and only if $g(y) \leq e$ whenever $y \in X$ with $f(y) = \lambda^{-1}$.*

Proof. First, fix $\lambda \in K \setminus \{\varepsilon\}$.

$\Longrightarrow$). Suppose that $g(x) \leq \tilde{f}(x)$ for all $x \in X$. Now, let $y \in X$ with $f(y) = \lambda^{-1}$ (note that $y \neq \perp$, because $\lambda \neq \varepsilon$ and f is topical). This together with the definition of $\tilde{f}$ implies that $e = \lambda\otimes f(y) = \lambda\dot{\otimes}f(y) = \tilde{f}(y)$. Therefore, by the hypothesis, one has $g(y) \leq \tilde{f}(y) = e$.

$\Longleftarrow$). Assume that for every $y \in X$ with $f(y) = \lambda^{-1}$, we have $g(y) \leq e$. Let $\widetilde{\ell_{y_1}}$ be an arbitrary minimal element of $\operatorname{supp}_u(\tilde{f}, \tilde{L})$ with $y_1 \neq \perp$. Then, by Proposition 4.1, $e = \tilde{f}(y_1) = \lambda\dot{\otimes}f(y_1) = \lambda \otimes f(y_1)$. This implies that $f(y_1) = \lambda^{-1}$ (it is worth noting that $f(y_1) \neq \top$. If $f(y_1) = \top$, then, $\tilde{f}(y_1) = \top$, which contradicts the fact that $\widetilde{\ell_{y_1}}$ is a minimal element of $\operatorname{supp}_u(\tilde{f}, \tilde{L})$ with $y_1 \neq \perp$). So, by the hypothesis, one has $g(y_1) \leq e$. Thus, in view of Proposition 2.1, $\widetilde{\ell_{y_1}} \in \operatorname{supp}_u(g, \tilde{L})$. This together with Proposition 4.2 and the fact that $g(y_1) \neq \varepsilon$ (by the hypothesis because $y_1 \neq \perp$) implies that there exists a minimal element $\widetilde{\ell_{y_2}}$ of $\operatorname{supp}_u(g, \tilde{L})$ such that

$$\widetilde{\ell_{y_2}}(x) \leq \widetilde{\ell_{y_1}}(x),\ \forall\ x \in X.$$

Therefore, we showed that for each minimal element $\widetilde{\ell_{y_1}}$ of $\operatorname{supp}_u(\tilde{f}, \tilde{L})$ with $y_1 \neq \perp$ there exists a minimal element $\widetilde{\ell_{y_2}}$ of $\operatorname{supp}_u(g, \tilde{L})$ such that $\widetilde{\ell_{y_2}}(x) \leq \widetilde{\ell_{y_1}}(x)$ for all $x \in X$. Thus, in view of Theorem 4.2 (the implication (2) $\Longrightarrow$ (1)), we have

$$\operatorname{supp}_u(\tilde{f}, \tilde{L}) \subseteq \operatorname{supp}_u(g, \tilde{L}). \tag{4.7}$$

Due to Proposition 4.3, the relation (4.7) implies that $g(x) \leq \tilde{f}(x)$ for all $x \in X$. Hence the proof is complete. □

In the sequel, let $f,\ g : X \longrightarrow \overline{K}$ be topical functions. Let

$$h(x) := g(x)\dot{\otimes}[f(x)]^{-1},\ \forall\, x \in X. \tag{4.8}$$

In the following, we give a necessary and sufficient condition for global maximum of the function h. In fact, we mean the function h as the difference of two topical functions f and g.

Theorem 4.4. *Suppose that the Assumptions (A0), (C1), (C2) and (C1) hold. Let $f, g : X \longrightarrow \overline{K}$ be strictly topical functions such that $f(x), g(x) \neq \varepsilon$ for all $x \in X \setminus \{\perp\}$. Let $\lambda \in K \setminus \{\varepsilon\}$ be such that $\inf_{x\in X} h(x) \geq \lambda$, where the function h is defined by (4.8). Then, $x_0 \in X$ is a global maximizer of the function h if and only if $g(y) \leq e$ whenever $y \in X$ with $f(y) = [h(x_0)]^{-1}$ (it is worth noting that $h(x_0) \in K \setminus \{\varepsilon\}$, because $\inf_{x\in X} h(x) \geq \lambda$ and $\lambda \in K \setminus \{\varepsilon\}$).*

In particular, if $h(x_0) = e$, then, x_0 is a global maximizer of the function h if and only if $\mathrm{supp}_u(f, \tilde{L}) \subseteq \mathrm{supp}_u(g, \tilde{L})$.

Proof. Define the function $\tilde{f} : X \longrightarrow \overline{K}$ by $\tilde{f}(x) := h(x_0)\dot{\otimes}f(x)$ for all $x \in X$. It is easy to see that $\tilde{f}$ is a strictly topical function and $\tilde{f}(x) \neq \varepsilon$ for all $x \in X \setminus \{\perp\}$.

Now, $x_0 \in X$ is a global maximizer of the function h, if and only if $h(x) \leq h(x_0)$ for all $x \in X$, if and only if

$$g(x) \leq h(x_0)\dot{\otimes}f(x) = \tilde{f}(x),\ \forall\, x \in X. \tag{4.9}$$

Therefore, in view of Theorem 4.3, one has the relation (4.9) is equivalent to $g(y) \leq e$ whenever $y \in X$ with $f(y) = [h(x_0)]^{-1}$. Now, assume that $h(x_0) = e$. Then, by the above, x_0 is a global maximizer of the function h if and only if $g(y) \leq e$ whenever $y \in X$ with $f(y) = e$, and so, by Theorem 4.3, if and only if $g(x) \leq f(x)$ for all $x \in X$, and by Theorem 4.2 (the implication (3) $\Longleftrightarrow$ (1)), if and only if $\mathrm{supp}_u(f, \tilde{L}) \subseteq \mathrm{supp}_u(g, \tilde{L})$, which completes the proof. □

5. Conclusions

We first characterized the superdifferential of extended valued topical functions. Next, we gave various characterizations of minimal elements of the upper support set of extended valued topical functions. As an application, we presented a necessary and sufficient condition for global maximum of the difference of two strictly topical functions defined on a semimodule with values in a semifield. These results have many applications in various parts of applied mathematics, mathematical economics and game theory, in particular, in the modelling of discrete event systems (see [11,12]).

Acknowledgements

The authors are very grateful to the anonymous referee for his/her useful suggestions on an earlier version of this paper. The comments of the referee were very fruitful and these comments have enabled the authors to improve the paper significantly.

References

[1] M. Akian, I. Singer, Topologies on lattice ordered groups, separation from closed downward sets and conjugations of type Lau, Optimization 52 (2003) 629–672.
[2] H. Barsam, H. Mohebi, Characterizations of upward and downward sets in semimodules by using topical functions, Numer. Funct. Anal. Optim. 37 (11) (2016) 1354–1377.
[3] G. Birkhoff, Lattice Theory, Colloquium Publications, vol. 25, Amer. Math. Soc., 1967.
[4] G.Y. Chen, X. Huang, X. Yang, Vector Optimization: Set-Valued and Variational Analysis, Lecture Notes in Economics and Mathematical Systems, vol. 541, Springer-Verlag, Berlin, 2005.
[5] G. Cohen, S. Gaubert, J.-P. Quadrat, I. Singer, Max-plus convex sets and functions, in: G.L. Litvinov, V.P. Maslov (Eds.), Idempotent Mathematics and Mathematical Physics, in: Contemp. Math., vol. 377, 2005, pp. 105–129.
[6] A.R. Doagooei, H. Mohebi, Optimization of the difference of ICR functions, Nonlinear Anal. 71 (2009) 4493–4499.
[7] A.R. Doagooei, H. Mohebi, Optimization of the difference of topical functions, J. Global Optim. 57 (2013) 1349–1358.
[8] S. Gaubert, J. Gunawardena, A Non-linear Hierarchy for Discrete Event Dynamical Systems, in: Proceedings of the 4th Workshop on Discrete Event Systems, Calgiari, Technical Report HPL-BRIMS-98-20, Hewlett-Packard Labs., 1998.
[9] S. Gaubert, S. Sergeev, Cyclic projectors and separation theorems in idempotent convex geometry, Fundam. Prikl. Math. 13 (4) (2007) 33–52, also e-print arXiv:0706.3347v1 [math.GM].
[10] A. Göpfert, H. Riahi, C. Tammer, C. Zălinescu, Variational Methods in Partially Ordered Spaces, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, vol. 17, Springer-Verlag, New York, 2003.
[11] J. Gunawardena, An Introduction to Idempotency, Cambridge University Press, Cambridge, 1998.
[12] J. Gunawardena, From Max-Plus Algebra to Non-expansive Mappings: A Non-linear Theory for Discrete Event Systems, Theoretical Computer Science, Technical Report HPL-BRIMS-99-07, Hewlett-Packard Labs., 1999.
[13] J. Gunawardena, M. Keane, On the Existence of Cycle Times for Some Non-expansive Maps, Technical Report HPL-BRIMS-95-003, Hewlett-Packard Labs., 1995.
[14] G.L. Litvinov, V.P. Maslov, G.B. Shpiz, Linear functionals and idempotent spaces, Dokl. Akad. Nauk 363 (3) (1998) 298–300, English translation: Dokl. Math. 58 (3) (1998) 389–391.
[15] G.L. Litvinov, V.P. Maslov, G.B. Shpiz, Idempotent functional analysis: an algebraic approach, Math. Zametki 69 (3) (2001) 758–797, English translation: Math. Notes 69 (5) (2001) 696–729.
[16] W.M. McEneaney, Max-Plus Methods for Nonlinear Control and Estimation, Birkhauser, Boston, 2006.
[17] H. Mohebi, Topical functions and their properties in a class of ordered Banach spaces, in: Continuous Optimization, Springer, 2005, pp. 343–360.
[18] H. Mohebi, S. Mirzadeh, Global minimization of the difference of increasing co-radiant and quasi-concave functions, in press.
[19] H. Mohebi, H. Momenaei, A study of downward sets with p-functions, Numer. Funct. Anal. Optim. 30 (3) (2009) 322–336.
[20] A.M. Rubinov, Abstract Convexity and Global Optimization, Kluwer Academic Publishers, Boston–Dordrecht–London, 2000.
[21] A.M. Rubinov, I. Singer, Topical and sub-topical functions, downward sets and abstract convexity, Optimization 50 (2001) 307–351.
[22] I. Singer, Abstract Convex Analysis, Wiley–Interscience, New York, 1997.
[23] I. Singer, On radiant sets, downward sets, topical functions and sub-topical functions in lattice ordered groups, Optimization 53 (2004) 393–428.

[24] I. Singer, Elementary topical functions on b-complete semimodules over b-complete idempotent semifields, Linear Algebra Appl. 433 (2010) 2139–2146.
[25] I. Singer, V. Nitica, Topical functions on semimodules and generalizations, Linear Algebra Appl. 437 (2012) 2471–2488.
[26] I. Singer, V. Nitica, Extended-valued topical and anti-topical functions on semimodules, Linear Algebra Appl. 446 (2014) 25–70.

5.2 Convergence of trajectories in infinite horizon optimization

The statistical convergence of a sequence of trajectories is studied in the first section of this paper. Then, we focus on the convergence of a sequence of minimizing trajectories in infinite horizon optimization problems. We consider the convergence of a sequence of trajectories in the sense of ideals as well as their particular case called the statistical convergence. Afterwards, we establish the I-convergence and statistical convergence of a sequence of minimizing trajectories.

Paper:

Convergence of trajectories in infinite horizon optimization

Accepted in:

International journal of nonlinear analysis and its applications

Authors: S. Hassani[a, b] and M. A. Mammadov[a, b]

[a]Federation University of Austalia, Victoria 3353, Australia

[b]National Information and Communications Technology Australia (NICTA)

Corresponding author:

Sara Hassani

e-mail: Sara.Hassani@nicta.com.au

Convergence of trajectories in infinite horizon optimization

S. Hassani,* and M.A. Mammadov

*Sara.Hassani@nicta.com.au

m.mammadov@federation.edu.au

Federation University of Austalia, Victoria 3353, Australia

National Information and Communications Technology Australia (NICTA)

Abstract

In this paper, we investigate the convergence of a sequence of minimizing trajectories in infinite horizon optimization problems. The convergence is considered in the sense of ideals and their particular case called the statistical convergence. The optimality is defined as a total cost over the infinite horizon.

Keywords. Infinite horizon optimization, Ideal convergence and statistical convergence.

AMS Subject Classification:90-XX

1 Introduction

Many important planning problems such as capacity expansion, equipment replacement and production planning involve sequences of related decisions over an infinite time horizon. The mathematical formulation of such problems lead to infinite horizon optimization which is the problem of selecting an infinite sequence of decisions such that the associated cost over an unbounded horizon is minimum [1, 2, 7, 10, 14, 16, 15, 17, 23].

In many studies an optimal solution/trajectory to infinite horizon problem is approximated by a sequence of finite horizon optimal solutions ([2, 18, 19, 23]). In [20], a general method for approximating optimal solution via the solutions to a simpler approximating problems is presented.

The uniqueness of optimal solution is a common assumption used in many studies [1, 2, 10]. For example in [2], under this assumption, an algorithm is developed for finding optimal solution and the results are applied to undiscounted Markov decision processes. Among the studies that do not use the uniqueness assumption we mention [17, 18, 23]. For example, in [18] a tie-breaking algorithm is presented based on selection of a nearest-point efficient solutions that converges to an optimal solution and the results are applied to the scheduling production problem to meet demand over an infinite horizon.

Since the cost over an unbounded horizon may be infinite or diverge, a discounting factor is applied in the definition of the total cost. It is clear that even in the presence of discounting, the total cost may still be infinite. In this case, different optimality criteria apart from minimal total cost are required [5, 12, 19, 21, 22]; the average cost [3, 8, 26], overtaking optimality[5, 13, 27] and 1-optimality [4, 25] are some examples of such optimality criteria.

In this paper, we consider systems described by the decision network as in [2]. These systems generate trajectories of decisions and there is a cost associated to each decision that could be used to define the functional - the total cost for a trajectory. The aim of this paper is to investigate the convergence of a sequence of trajectories under the assumption that the functional values (total costs) converge to the optimal value (i.e. the minimal total cost). The convergence is considered in the sense of ideals and their particular case called the statistical convergence.

The paper is organized as follows. Notations and the problem statement are presented in the next section. Some preliminary results about convergence of the sequence of trajectories are established in Section 3. The I-convergence and the statistical convergence of a sequences of trajectories are considered in Section 4.

2 Notations and problem statement

We begin with the decision network, (Σ, A, C), where Σ is the set of states (nodes), A is the set of decisions (arcs) and C is a real-valued cost function $C : A \to R$. We assume that the decision network satisfies the following conditions [2]:

- there is a node called single root with the following properties
  - there is no incoming arcs to this node,
  - every other node can be reached from the single root,
- the set of decisions available at any node is nonempty and finite,
- the set of incoming decisions to any node is also finite.

Under these assumptions, it has been proved that [24, Theorem 1] the set of nodes can be numbered as $\Sigma = \{\sigma_1, \sigma_2, \sigma_3, ...\}$ such that if $(\sigma_i, \sigma_j) \in A$ where $\sigma_i, \sigma_j \in N$, then $i < j$.

An infinite trajectory $\mathbf{s}$ is an infinite sequence of states $(s_1, s_2, s_3, ...)$ where s_1 is a given fixed root, $s_i \in \Sigma$ and $(s_i, s_{i+1}) \in A$ for all $i = 1, 2, ...$. The cost $C(s_i, s_{i+1})$ associates with the decision (s_i, s_{i+1}). The set of all trajectories $\mathbf{s}$ is denoted by $\prod$.

Now we introduce the metric in the set of trajectories. Consider two trajectories $\mathbf{s} = (s_1, s_2, s_3, ...)$ and $\mathbf{s}' = (s'_1, s'_2, s'_3, ...)$. In [2], the metric ρ, on $\prod$ is constructed as follows:

$$\rho(\mathbf{s}, \mathbf{s}') = \sum_{i=1}^{\infty} \phi_i(\mathbf{s}, \mathbf{s}') 2^{-i}, \tag{2.1}$$

where

$$\phi_i(\mathbf{s}, \mathbf{s}') = \begin{cases} 0 & \text{if} \quad s_i = s'_i \\ 1 & \text{otherwise} \end{cases} .$$

In [2] (Lemma 1), it is proved that the set $\prod$ is complete and hence compact in the sense of this metric.

Under this metric, the closeness of trajectories depends on the number of initial nodes over which they agree. For example, given any $i \in \mathbf{N}$ it can easily be verified that the following holds:

$$\rho(\mathbf{s}, \mathbf{s}') > \frac{1}{2^i} \quad \Rightarrow \quad s'_r \neq s_r, \ \exists r \in \{1, 2, \cdots, i\}. \tag{2.2}$$

Functional - the total cost $f(\mathbf{s})$ of trajectory $\mathbf{s}$ is defined as in [2] given by

$$f(\mathbf{s}) = \sum_{i=1}^{\infty} C(s_i, s_{i+1}). \tag{2.3}$$

We will assume that f is uniformly convergent over $\prod$; that is, for any $\varepsilon > 0$ there exists n_ε such that for all trajectories $\mathbf{s}$ the relation $\sum_{i=n}^{\infty} C(s_i, s_{i+1}) < \varepsilon$ holds for all $n \geq n_\varepsilon$. In this case f is continuous on $\prod$. Note that this is not a restrictive assumption; it holds if the cost function $C(s_i, s_{i+1})$ is uniformly bounded and also is discounted, for example, by $(1/2)^i$ (see Assumption 1 and Lemma 2 in [2]).

We consider the following optimization problem

$$\text{Minimize} \ \ f(\mathbf{s}), \quad \text{subject to} \quad \mathbf{s} \in \prod. \tag{2.4}$$

Since f is continuous and $\prod$ is compact, an optimal solution s^* to problem (2.4) exists. We call $\mathbf{s}^n$ a minimizing sequence if $f(\mathbf{s}^n)$ converges to the minimal value $f(\mathbf{s}^*)$ of the objective function in this problem. The aim of this paper is to investigate the convergence of minimizing sequence $\mathbf{s}^n$ to $\mathbf{s}^*$ by considering different types of convergence.

3 Preliminary results

In this section, we consider the convergence of a sequence of trajectories $\{\mathbf{s}^n\}_{n\in\mathbf{N}}$ to the trajectory $\mathbf{s}$ in the sense of ideals as well as their particular case called the statistical convergence. We recall that the initial point of all sequences is the same; that is, $s_1 = s_1^n$ for all n. We will use the notation $\{\{\mathbf{s}\}\} := \{s_1, s_2, s_3, ...\}$ to denote the set of states for trajectory $\mathbf{s}$.

First we give the definition of ideal and I-convergence.

Definition 3.1. A family $I \subset 2^{\mathbf{X}}$ of subsets of a nonempty set $\mathbf{X}$ is said to be an ideal in X if

- $A, B \in I$ implies $A \cup B \in I$,
- $B \subset A,\ A \in I$ implies $B \in I$,

while an admissible ideal I of $\mathbf{X}$ further satisfy $\{x\} \in I$ for each $x \in \mathbf{X}$.

Clearly, an ideal admissible contains all finite sets in $\mathbf{X}$. In the remainder of this section, we assume that any ideal is admissible and I is an ideal in $\mathbf{N}$.

Definition 3.2. [6, 11] A sequence $\mathbf{s}^n$ in a metric space $(\mathbf{X}, \rho)$ is said to be I-convergent to $\mathbf{s} \in \mathbf{X}$ (in short $\mathbf{s} = \mathrm{I} - \lim_{n\to\infty} \mathbf{s}^n$) if $K(\epsilon) \in I$ for each $\epsilon > 0$, where $K(\epsilon) = \{n \in \mathbf{N} :\ \rho(\mathbf{s}^n, \mathbf{s}) \geq \epsilon\}$.

Below we consider two special cases of ideals.

1. Classical convergence. In this case the ideal is the set of all finite subsets of $\mathbf{N}$; that is

$$I = I_{fin} \doteq \{A \subset \mathbf{N} :\ |A| < \infty\}.$$

Clearly, both of the conditions in Definition 3.1 are satisfied.

2. Statistical convergence. First we give the definition of the statistical convergence in terms of the notion of density. Assume K is a subset of the positive integers $\mathbf{N}$. $K_n = \{k \in K : \ k \leq n\}$ and $|K_n|$ denotes the number of elements in K_n. The natural density of K is given by $\delta(K) = \lim_{n\to\infty} \frac{|K_n|}{n}$. It may not exist; in this case the upper and lower asymptotic densities for the set K are defined as follows:

$$\bar{\delta}(K) = \limsup_{n\to\infty} \frac{|K_n|}{n} \text{ and } \underline{\delta}(K) = \liminf_{n\to\infty} \frac{|K_n|}{n}.$$

Note that $\underline{\delta}(K) \leq \delta(K) \leq \bar{\delta}(K)$.

Definition 3.3. [9] A sequence $\{\mathbf{s}^n\}_{n\in\mathbf{N}}$ is statistically convergent to $\mathbf{s}$ provided that for every $\epsilon > 0$, the set $K(\epsilon) = \{n \in \mathbf{N} : \ \rho(\mathbf{s}^n, \mathbf{s}) \geq \epsilon\}$ has natural density zero.

It is not difficult to observe that both of the conditions in Definition 3.1 are satisfied if we define the ideal as the set of subsets of $\mathbf{N}$ with density zero. Thus in this case we set

$$I = I_d \doteq \{A \subset \mathbf{N} : \ \delta(A) = 0\}.$$

In the next lemma, the convergence of the sequence $\mathbf{s}^n$ to the trajectory $\mathbf{s}$ is considered.

Lemma 3.4. *Assume that* $\delta(\{n \in \mathbf{N} : \ s_i^n \neq s_i\}) = 0$ *for all* $i \in \mathbf{N}$. *Then sequence* $\mathbf{s}^n$ *statistically converges to* $\mathbf{s}$.

Proof: Take an arbitrary $\epsilon > 0$ and denote $A_\epsilon := \{n \in \mathbf{N} : \ \rho(\mathbf{s}^n, \mathbf{s}) > \epsilon\}$. We show that $\delta(A_\epsilon) = 0$.

Let $r_\epsilon \in \mathbf{N}$ such that $\frac{1}{2^{r_\epsilon}} < \epsilon$. Consider the sets

$$A_{r_\epsilon} := \{n \in \mathbf{N} : \ \rho(\mathbf{s}^n, \mathbf{s}) > \frac{1}{2^{r_\epsilon}}\}$$

and

$$B_{r_\epsilon} := \{n \in \mathbf{N} : \ s_i^n \neq s_i^* \text{ for some } i \in \{1, 2, ..., r_\epsilon\}\}.$$

From (2.2) we have $A_\epsilon \subset A_{r_\epsilon} \subset B_{r_\epsilon}$. On the other hand, B_{r_ϵ} can be represented in the form

$$B_{r_\epsilon} = \cup_{i=1}^{r_\epsilon} \{n \in \mathbf{N} : \ s_i^n \neq s_i\}.$$

By the assumption of the lemma, $\delta(\{n \in \mathbf{N} : \ s_i^n \neq s_i\}) = 0$ for all $i = 1, ..., r_\epsilon$ and therefore $\delta(B_{r_\epsilon}) = 0$. Thus, since $A_\epsilon \subset A_{r_\epsilon} \subset B_{r_\epsilon}$ we have $\delta(A_\epsilon) = 0$.

Lemma is proved. □

It is clear that under the conditions of lemma 3.4, the classical convergence may not be true. Indeed, for example, if $\mathbf{s}^n = \mathbf{s}$ for all $n \in \mathbf{N} \setminus \{3^j\}_{j \in \mathbf{N}}$ and $\mathbf{s}^n = (s_1, \bar{s_2}, s_3, s_4, s_5, ...)$ for $n \in \{3^j\}_{j \in \mathbf{N}}$ where $\bar{s_2} \neq s_2$ then it is not difficult to show that $\mathbf{s}^n$ is statistically convergent to $\mathbf{s}$ while the classical convergence is not true.

4 Convergence of a sequence of minimizing trajectories

In this section, we investigate the convergence of the minimizing sequence $\mathbf{s}^n$ to the optimal trajectory $\mathbf{s}^*$ of the problem (2.4); that is, under the assumption that $f(\mathbf{s}^n) \to f(\mathbf{s}^*)$ we investigate the convergence $\mathbf{s}^n \to s^*$ as $n \to \infty$. We do not assume the uniqueness of $\mathbf{s}^*$; however, we will consider a fixed optimal trajectory $\mathbf{s}^*$ and will formulate the main assumptions by using this trajectory. Note that $\mathbf{s}^n$ may not converge to $\mathbf{s}^*$, in this case the $I-$convergence and statistical convergence will be considered.

Given sequence $\mathbf{s}^n$ and set $K \subset \mathbf{N}$ we define

$$H(K) = \{j \in \mathbf{N} : \ s_j^* \in \{\{\mathbf{s}^n\}\}, \ \forall \ n \in K\}. \tag{4.1}$$

In the case $K = \mathbf{N}$ for the sake of simplicity we denote

$$H = H(\mathbf{N}) = \{j \in \mathbf{N} : \ s_j^* \in \{\{\mathbf{s}^n\}\}, \ \forall \ n \in \mathbf{N}\}. \tag{4.2}$$

For trajectory $\mathbf{s}$, we denote the section connecting two nodes a and b by

$$P(\mathbf{s} : a, b) = \{s_{n_1}, s_{n_1+1}, ..., s_{n_2-1}, s_{n_2}\};$$

where $s_{n_1} = a$ and $s_{n_2} = b$.

The corresponding cost is

$$f(P(\mathbf{s} : a, b)) = C(s_{n_1}, s_{n_1+1}) + ... + C(s_{n_2-1}, s_{n_2}).$$

Let $\mathbf{s}^* = (s_1^*, s_2^*, s_3^*, s_4^*, ...)$.

Condition A: For any $s_i^*, s_j^* \in \mathbf{s}^*$ with $i < j$ and any trajectory connecting points s_i^*, s_j^* (that is, $s_{n_i} = s_i^*, s_{n_j} = s_j^*$ and $P(\mathbf{s} : s_{n_i}, s_{n_j}) \neq P(\mathbf{s}^* : s_i^*, s_j^*)$) the following inequality holds

$$f(P(\mathbf{s} : s_{n_i}, s_{n_j})) - f(P(\mathbf{s}^* : s_i^*, s_j^*)) > 0.$$

A trajectory is said to be efficient if it reaches each of the states through which it passes at minimum cost. Efficient solutions are shown in [19] to be average-cost optimal under a state reachability property. Clearly, Condition A implies that $\mathbf{s}^*$ is an efficient trajectory.

Condition A means any finite section $(s_i^*, s_{i+1}^*, ..., s_j^*)$ in terms of an "optimality" connecting nodes s_i^* and s_j^* is unique. Note also that as $s_i^*, s_j^* \in \Sigma$ and there is a finite number of nodes between s_i^* and s_j^*; that is, the number of trajectories connecting these two nodes is finite.

In the following example, we show that Condition A does not mean the uniqueness of optimal trajectory.

Example 4.1. Assume that the set of states (nodes) in a decision network consists of nodes $\{s_k\}_{k \in \mathbf{N}}$ and the cost function is given by

$$C(s_1, s_2) = \frac{1}{2}, \ C(s_{2k-1}, s_{2k+1}) = \frac{1}{2^k}, \ C(s_{2k}, s_{2k+1}) = C(s_{2k}, s_{2k+2}) = \frac{1}{2^{k+1}}, \ \ \forall k \geq 1. \tag{4.3}$$

Clearly, $\mathbf{s}^* = (s_1, s_3, s_5, s_7, \cdots)$ is an optimal trajectory. We construct a sequence of trajectories $\mathbf{s}^n$ in the following form:

$$\mathbf{s}^n = (s_1, s_2, s_{2+2}, \cdots, s_{2n}, s_{2n+1}, s_{2n+3}, \cdots), \quad n \in \mathbf{N}. \tag{4.4}$$

We show that the condition A holds. For this aim, it is enough to show that condition A holds for $s_1, s_{2j+1} \in \mathbf{s}^*, \ j \geq 1$. Taking into account the definition of the cost function in (4.3), we have

$$C(s_1, s_2)+(\sum_{k=1}^{j-1} C(s_{2k}, s_{2k+2}))+C(s_{2j}, s_{2j+1})+\sum_{k=j}^{\infty} C(s_{2k-1}, s_{2k+1}) = \frac{1}{2^{j+1}} > 0;$$

that is, condition A holds.

On the other hand, for the trajectory $\bar{\mathbf{s}} = (s_1, s_2, s_4, ..., s_{2n}, s_{2n+2}, ...)$ the relation $f(\bar{\mathbf{s}}) = f(\mathbf{s}^*)$ holds which means optimal trajectory $\mathbf{s}^*$ is not unique. $\square$

In the next theorem, we establish the I-convergence of the sequence $\mathbf{s}^n$ to the optimal trajectory $\mathbf{s}^*$ when $f(\mathbf{s}^n) \to f(\mathbf{s}^*)$.

Theorem 4.2. *Assume that optimal trajectory $\mathbf{s}^*$ satisfies Condition A, $\mathbf{s}^n$ is a minimizing sequence and there exists $K \subset \mathbf{N}$ such that $|H(K)| = \infty$ and $|K \cap A| = \infty$ for all $A \notin I$. Then $\mathbf{s}^n$ is I-convergent to $\mathbf{s}^*$ as $n \to \infty$.*

Proof: On the contrary assume there exists $\epsilon > 0$ such that $\{n \in \mathbf{N} : \rho(\mathbf{s}^n, \mathbf{s}^*) > \epsilon\} \notin I$. Take any $r_\varepsilon \in \mathbf{N}$ satisfying $\frac{1}{2^{r_\varepsilon}} < \epsilon$ and denote

$$A_{r_\varepsilon} = \{n \in \mathbf{N} :\ \rho(\mathbf{s}^n, \mathbf{s}^*) > \frac{1}{2^{r_\varepsilon}}\}. \tag{4.5}$$

Clearly $\{n \in \mathbf{N} :\ \rho(\mathbf{s}^n, \mathbf{s}^*) > \epsilon\} \subset A_{r_\varepsilon}$ and therefore $A_{r_\varepsilon} \notin I$. From (2.2) it follows that

$$A_{r_\varepsilon} \subset \{n \in \mathbf{N} :\ s_i^n \neq s_i^*,\ \ \exists\, i \in \{2, 3, ..., i\}\}.$$

By the assumption of the lemma we have $|K \cap A_{r_\varepsilon}| = \infty$. Consider the set $H(K)$ defined in (4.1) and denote

$$t = \min\{m :\ r_\varepsilon \leq m,\ m \in H(K)\ \}.$$

We note that such t exists as $|H(K)| = \infty$.

Now for any $n \in K \cap A_{r_\varepsilon}$ the relation $s_t^* \in \{\{\mathbf{s}^n\}\}$ holds; that is, $s_{j_n}^n = s_t^*$ for some index j_n. Denote

$$\alpha = \inf_{n \in K \cap A_{r_\varepsilon}} \{\sum_{r=1}^{j_n - 1} C(s_r^n, s_{r+1}^n) - \sum_{j=1}^{t-1} C(s_j^*, s_{j+1}^*)\}. \tag{4.6}$$

We note that there are only a finite number of possible different combinations $(s_1^n, \cdots, s_{j_n}^n)$ with the same fixed initial point s_1^* and the end point s_t^*. Then, condition A implies $\alpha > 0$.

Denote

$$a^n = \sum_{r=1}^{j_n - 1} C(s_r^n, s_{r+1}^n),\ \ a = \sum_{j=1}^{t-1} C(s_j^*, s_{j+1}^*);$$

$$b^n = \sum_{r=j_n}^{\infty} C(s_r^n, s_{r+1}^n),\ \ b = \sum_{j=t}^{\infty} C(s_j^*, s_{j+1}^*).$$

Clearly, $f(\mathbf{s}^n) = a^n + b^n$ and $f(\mathbf{s}^*) = a + b$. From (4.6) we have $a^n \geq a + \alpha$. Since $\mathbf{s}^*$ is optimal and $s_{j_n}^n = s_t^*$ we have $b^n \geq b$. Thus

$$f(\mathbf{s}^n) = a^n + b^n \geq a^n + b \geq a + b + \alpha = f(\mathbf{s}^*) + \alpha,\ \ \forall n \in K \cap A_{r_\varepsilon}.$$

This means that $f(\mathbf{s}^n)$ does not converge to $f(\mathbf{s}^*)$; that is, $\mathbf{s}^n$ is not a minimizing sequence. This is a contradiction. Theorem is proved.$\square$

In this lemma, condition $|K \cap A| = \infty$ for all $A \notin I$ means that the set K should be quite "large". We describe it in Corollary 4.5 in terms of the density of K.

In the next lemma, we investigate the classical convergence of the sequence $\{\mathbf{s}^n\}_{n\in\mathbf{N}}$ to the optimal trajectory $\mathbf{s}^*$. It is shown that stronger condition is required in comparison to Theorem 4.2.

Corollary 4.3. *Assume that optimal trajectory* $\mathbf{s}^*$ *satisfies Condition A,* $\mathbf{s}^n$ *is a minimizing sequence and* $|H| = \infty$. *Then* $\mathbf{s}^n \to \mathbf{s}^*$ *as* $n \to \infty$.

Here H is defined in (4.2) which corresponds to $\mathbf{K} = \mathbf{N}$ in terms of Theorem 4.2.

Proof: We apply Theorem 4.2 assuming that $\mathbf{K} = \mathbf{N}$ and the ideal I is the set of finite subsets of $\mathbf{N}$; that is, $I = I_{fin}$. Firstly, we have $|H(\mathbf{K})| = |H| = \infty$. On the other hand, for any $A \notin I_{fin}$ the relation $|A| = \infty$ holds and therefore

$$|\mathbf{K} \cap A| = |\mathbf{N} \cap A| = |A| = \infty.$$

Thus, all the assumptions of Theorem 4.2 hold. The ideal convergence in this case is the classical convergence $\mathbf{s}^n \to \mathbf{s}^*$ as $n \to \infty$.

The corollary is proved. $\square$

The condition $|\mathbf{H}| = \infty$ means the number of nodes in $\mathbf{s}^*$ that is "common" in all trajectories $\mathbf{s}^n$ is infinite; in other words, all trajectories $\mathbf{s}^n$ pass through an infinite number of nodes in $\mathbf{s}^*$.

In the following example, we investigate the necessity of condition A in this corollary.

Example 4.4. Let $\mathbf{s}^* = (s_1, s_3, s_4, \cdots)$, $\mathbf{s}^n = (s_1, s_2, s_3, s_4, \cdots)$ and

$$C(s_1, s_2) + C(s_2, s_3) = C(s_1, s_3).$$

Then $f(\mathbf{s^n}) = f(\mathbf{s}^*)$ for all n; however $\mathbf{s^n}$ does not converge to $\mathbf{s}^*$ as $\rho(\mathbf{s^n}, s^*) \geq 0.5$ for all n. We also mention that in this example, $H = \mathbf{N}$. $\square$

To consider the necessity of condition $|H| = \infty$ in Corollary 4.3, we refer to the decision network in Example 4.1. In this example, given any $k \geq 2$, the relation $s_{2k-1} \notin \{\{\mathbf{s}^n\}\}$ holds for all $n \geq k$. This means that the set H contains just one element; that is, the condition $|H| \neq \infty$ does not hold. Clearly, $f(\mathbf{s^n}) \to f(\mathbf{s}^*)$ however $\mathbf{s}^n$ does not converge to $\mathbf{s}^*$.

Now we consider a special case of Theorem 4.2 when the ideal convergence is defined by the statistical convergence. We have

Corollary 4.5. *Assume that optimal trajectory* $\mathbf{s}^*$ *satisfies Condition A,* $\mathbf{s}^n$ *is a minimizing sequence, there exists* $K \subset \mathbf{N}$ *such that* $|H(K)| = \infty$ *and* $\delta(K) = 1$. *Then* $\mathbf{s}^n$ *statistically converges to* $\mathbf{s}^*$ *as* $n \to \infty$.

Proof: We apply Theorem 4.2 assuming that the ideal I is the set of subsets of $\mathbf{N}$ having density 0; that is, $I = \{A \subset \mathbf{N} : \delta(A) = 0\}$. The relation $A \notin I$ in this case means A has a nonzero density. Then for any set $\mathbf{K} \subset \mathbf{N}$ with $\delta(\mathbf{K}) = 1$, $\mathbf{K} \cap A$ also has nonzero density. This means that $|\mathbf{K} \cap A| = \infty$ and all the assumptions of Theorem 4.2 hold. The corollary is proved. $\square$

Clearly, Corollary 4.5 is a special case of Corollary 4.3. Next we provide an example where $\mathbf{s}^n$ statistically converges to $\mathbf{s}^*$ however the classical convergence $\mathbf{s}^n \to \mathbf{s}^*$ is not true.

Example 4.6. Assume that the set of states (nodes) in a decision network consists of nodes $\{s_k\}_{k\in\mathbf{N}}$, $\{\xi_k\}_{k\in\mathbf{N}}$ and the cost function is given by

$$C(s_1, s_2) = \frac{1}{2},\ C(s_{2k-1}, s_{2k+1}) = \frac{1}{2^k},\ C(s_{2k}, s_{2k+1}) = C(s_{2k}, s_{2k+2}) = \frac{1}{2^{k+1}},\ \ \forall k \geq 1. \tag{4.7}$$

$$C(s_{3^{\kappa(n)-1}}, \xi_n) = C(\xi_n, s_{3^{\kappa(n)-1}+2}) = \frac{1}{2}\left[C(s_{3^{\kappa(n)-1}}, s_{3^{\kappa(n)-1}+2}) + \frac{1}{n}\right],\ \ \forall n \in \mathbf{N}\backslash\{3^k\}_{k\in\mathbf{N}}. \tag{4.8}$$

In this example $\mathbf{s}^* = (s_1, s_3, s_5, s_7, \cdots)$ is an optimal trajectory. Consider the function of indices $\kappa : \mathbf{N} \to \mathbf{N}$ defined by

$$\kappa(n) = i,\ \forall n \in \{3^{i-1}, 3^{i-1}+1, \cdots, 3^i - 1\},\ i = 1, 2, \cdots.$$

We construct a sequence of trajectories s^n in the following form:

$$\mathbf{s}^n = (s_1, s_2, s_{2+2}, \cdots, s_{2n}, s_{2n+1}, s_{2n+3}, \cdots),\quad n \in \{3^k\}_{k\in\mathbf{N}};$$

$$\mathbf{s}^n = (s_1, s_3, \cdots, s_{3^{\kappa(n)-1}}, \xi_n, s_{3^{\kappa(n)-1}+2}, s_{3^{\kappa(n)-1}+4}, \cdots),\quad n \in \mathbf{N} \setminus \{3^k\}_{k\in\mathbf{N}}.$$

Now we show that $f(\mathbf{s^n}) \to f(\mathbf{s}^*)$. For any $n \in \{3^k\}_{k\in\mathbf{N}}$, from (4.7) we have

$$f(\mathbf{s^n}) - f(\mathbf{s}^*) = \frac{1}{2^{n+1}}.$$

Similarly, for any $n \in \mathbf{N} \setminus \{3^k\}_{k\in\mathbf{N}}$, it follows from (4.8) that

$$f(\mathbf{s}^n) - f(\mathbf{s}^*) = \frac{1}{n}.$$

Therefore $f(\mathbf{s^n}) \to f(\mathbf{s}^*)$ as $n \to \infty$; that is, s^n is a minimizing sequence.

Now consider the set $K = N \setminus \{3^k\}_{k\in\mathbf{N}}$. Clearly, $H(K) = \{2n-1\}_{n\in\mathbf{N}}$ and $\delta(K) = 1$; that is the conditions of Corollary 4.5 are satisfied. It is not difficult to verify that $\mathbf{s}^n$ statistically converges to $\mathbf{s}^*$. However, it does not converge to $\mathbf{s}^*$ in the sense of classical convergence as $\rho(\mathbf{s^n}, s^*) \geq 0.5$ for all $n \in \{3^k\}_{k\in\mathbf{N}}$. $\square$

References

[1] J. Bean and R. Smith. Conditions for the existence of planning horizons. *Math. Opns. Res*, (9):391–401, 1984.

[2] J. C. Bean and R. L. Smith. Conditions for the discovery of solution horizons. *Mathematical programming*, 59:215–229, 1993.

[3] D. P. Bertsekas. *Dynamic Programming and Stochastic Control*. Academic Press, New-York, 1976.

[4] D. Blackweil. Discrete dynamic programming. *Ann.Math.Statist.*, 33:719–726, 1962.

[5] D. A. Carlson, A. Haurie, and A. Leizarowitz. *Infinite Horizon Optimal Control: Deterministic and Stochastic Systems.* 2nd, rev. and enl. ed. Berlin: Springer-Verlag, 1991.

[6] H. Cartan. Filtres et ultrafiltres. *C. R. Acad. Sci. Paris*, 205:777–779, 1937.

[7] R. Dorfman, P. Samuelson, and R. Solow. *Linear programming and economic analysis*. New York: McGraw-Hill, 1958.

[8] V. Gaitsgory and S. Rossomakhine. Linear programming approach to deterministic long run average problems of optimal control. *SIAM journal on control and optimization*, 44(6):2006–2037, 2006.

[9] H.Fast. Sur la convergence statistique. *Colloq. Math*, 2:241–244, 1951.

[10] W. Hopp, J. Bean, and R. Smith. A new optimality criterion for non-homegeneous markov decesion process. *Opns. Res*, (35):875–883, 1987.

[11] P. Kostyrko, T. *Šalát*, and W. Wilczyński. I-convergence. *Real Analysis Exchange*, 26(2):669–685, 2000/2001.

[12] J. B. Lasserre. Decision horizon, overtaking and 1-optimality criteria in optimal control. *Advances in Optimization. Lecture Notes in Mathematics, no. 302, Eiselt, H. A. and G. Pederzoli, eds. Springer Verlag, New York*, pages 247–261, 1988.

[13] A. Leizarowitz. Infinite horizon autonomous systems with unbounded cost. *Appl Math Optim*, 13:19–43, 1985.

[14] V. Makarov and A. Rubinov. *Mathematical theory of economic dynamics and equilibria*. Springer-Verlag, New York, 1977.

[15] L. W. McKenzie. Turnpike theory. *Econometrica*, 44:841–866, 1976.

[16] S. Rayan. *Degenercy in discrete infinite horizon optimization.* PhD thesis, Department of industrial and operations engineering, The University of Michigan, Ann Arbor, 1988.

[17] S. M. Rayan, J. C. Bean, and R. L. Smith. A tie-breaking rule for discrete infinite horizon optimization. *Operation Research*, 40(1):117–126, 1992.

[18] I. Schochetman and R. Smith. *Finite dimensional approximation in infinite dimensional mathematical programming, Mathematical Programming*, 54:307–333, 1992.

[19] I. Schochetman and R. Smith. Existence and discovery of average optimal solutions in deterministic infinite horizon optimization. *Mathematics of Operations Research*, pages 416–432, 1998.

[20] I. Schochetman and R. Smith. A finite algorithm for solving infinite dimensional optimization problems. *Annals of Operations Research*, 101:119142, 2001.

[21] I. Schochetman and R. Smith. Optimality criteria for deterministic discrete-time infinite horizon optimization. *International Journal of Mathematics and Mathematical Sciences*, pages 57–80, 2004.

[22] I. Schochetman and R. Smith. Existence of efficient solutions in infinite horizon optimization under continuous and discrete controls. *Operations Research Letters*, (33):97–104, 2005.

[23] I. E. Schochetman and R. L. Smith. Infinite horizon optimization. *Mathematics of Operations Research*, 14(3):559–574, 1989.

[24] D. Skilton. Imbedding posets in the integers. *Order*, 1(3):229–233, 1985.

[25] A. F. Veinott. On finding optimal policies in discrete dynamic programming with no discounting. *Ann. Math. Statist.*, 37(5):1284–1294, 1966.

[26] A. O. Wachs, I. E. Schochetman, and R. L. Smith. Average optimality in nonhomogeneous infinite horizon markov decision processes. *Mathematics of Operations Research*, 36(1):147–164, 2011.

[27] A. Zaslavski. *Turnpike properties in the calculus of variations and optimal control.* Springer, 2006.

5.3 Optimality conditions in infinite horizon optimization by contingent derivative

In this section, we consider optimization problems where objective function is the total cost over an infinite trajectory. We first define new notions contingent cone and upper derivative contingent cone. Then, we derive optimality conditions for this class of optimization problems by using newly defined notions.

Paper:

Optimality conditions in infinite horizon optimization by contingent derivative

Submitted to:

Pacific Journal of Optimization

Authors: S.Hassani[a, b] and M.A.Mammadov[a]

[a]Federation University Australia, Victoria 3353, Australia

[b]National Information and Communications Technology Australia (NICTA)

Corresponding author:

Sara Hassani

e-mail: sarahassani@students.federation.edu.au

Optimality conditions in infinite horizon optimization by contingent derivative

S. Hassani *[1,2], and M.A. Mammadov [1]

sara.hassani@nicta.com.au

m.mammadov@federation.edu.au

[1] *Federation University of Austalia, Victoria 3353, Australia*

[2] *National Information and Communications Technology Australia (NICTA)*

Abstract

In this paper, the notion of contingent cone to the set of trajectories in infinite horizon optimization problems is introduced. Some important properties of contingent cone are investigated. Also the notion of upper contingent derivative is introduced that is based on the same idea used in the literature when defining with directional derivatives. Then, optimality conditions are derived in terms of the contingent cone and the upper contingent derivative.

Keywords. Infinite horizon optimization, optimal trajectory, contigent cone, upper contingent derivative, optimality condition.

AMS Subject Classification: 90C26

1 Introduction

Infinite horizon optimizations are an important class of optimization problems where the objective function is often defined as a total cost over infinite horizon [1, 2, 8, 14, 17, 19, 18, 20, 25]. This class has many applications in inventory control, production planning, equipment replacement and capacity expansion.

The total cost over an unbounded horizon may be infinite or diverge. Taking this factor into account, different optimality criteria apart from minimal total have been considered [6, 15, 22, 23, 24]. For example, a discounting factor is used to guarantee the convergence of a sum of infinitely many costs over an infinite horizon [1, 2]. Some other examples of such optimality criteria are efficiency or finite optimality [13, 21, 26], the average cost [3, 11, 29], overtaking optimality [6, 12, 16, 30, 31] and 1-optimality [4, 28].

The notion of tangent cone (contingent cone) plays an important role in driving optimality conditions. The tangent cone gives an approximation of a set around given point. Different definitions are introduced in the literature for a tangent cone, such as Bouligand tangent cone [5] and Clarke tangent cone [7]. The use of tangent cones in optimization was initiated by Dubovitskii and Miljutin [9, 10].

In this paper, we consider systems described by the decision network as in [2]. These systems generate trajectories of decisions and there is a cost associated to each decision that could be used to define the functional - total cost for a trajectory. The aim of this paper is to introduce a contingent cone and directional derivative for metric space of trajectories and investigate optimality conditions.

The paper is organized as follows. Notations, problem statement and the new notion “upper contingent derivative” are presented in the next section. “Contingent cone” and its some important properties are established in Section 3. Optimality conditions by applying new introduced notions are considered in 4.

2 Notations and problem statement

Consider the decision network, (Σ, A, C), where Σ is the set of states (nodes), A is the set of decisions (arcs) and C is a real-valued cost function $C : A \to R$. The notation $\mathbb{N}$ is used to denote the set of natural numbers $\{1, 2, \cdots\}$. Throughout the paper we will assume that the following conditions hold:

1. there is a node called single root with the following properties
   - there is no incoming arcs to this node,
   - every other node can be reached from the single root,
2. the set of decisions available at any node is nonempty and finite,
3. the set of incoming decisions to any node is also finite.

It has been proved that [27, Theorem 1] under these assumptions, the set of nodes can be numbered, say as in the form $\Sigma = \{\sigma_1, \sigma_2, \sigma_3, ...\}$, such that the following holds: if $(\sigma_i, \sigma_j) \in A$ for some nodes $\sigma_i, \sigma_j \in \Sigma$, then $i < j$.

Definition 2.1. A trajectory $\mathbf{s}$ is an infinite sequence of states $(s_1, s_2, s_3, ...)$ where $s_1 = \sigma_1$ is a given fixed root, $s_i \in \Sigma$ and $(s_i, s_{i+1}) \in A$ for all $i = 1, 2, ...$.

The set of all trajectories $\mathbf{s}$ will be denoted by Π. This set can be endowed by a metric. We will use the metric used in [2]; namely, given any two trajectories $\mathbf{s} = (s_1, s_2, s_3, ...)$ and $\mathbf{s}' = (s'_1, s'_2, s'_3, ...)$. the metric ρ is defined as follows:

$$\rho(\mathbf{s}, \mathbf{s}') = \sum_{i=1}^{\infty} \phi_i(\mathbf{s}, \mathbf{s}') 2^{-i}, \tag{2.1}$$

where

$$\phi_i(\mathbf{s}, \mathbf{s}') = \begin{cases} 0 & \text{if} \quad s_i = s'_i \\ 1 & \text{otherwise} \end{cases}.$$

In [[2]; Lemma 1], it is proved that the set Π is complete and hence compact in the sense of this metric.

Under this metric, the closeness of trajectories depends on the number of initial nodes over which they agree. For example, given any $i \in \mathbb{N}$, it can easily be verified that the following holds:

$$\rho(\mathbf{s}, \mathbf{s}') \leq \frac{1}{2^{k+1}} \quad \Rightarrow \quad s'_i = s_i, \ \forall i = 1, 2, \cdots, k. \tag{2.2}$$

Optimization problem. Given trajectory $\mathbf{s} = (s_1, s_2, s_3, ...)$, the value $C(s_i, s_{i+1})$ is the cost associated with the decision (s_i, s_{i+1}). Then, the objective function in this problem can be determined as the total cost over the trajectory $\mathbf{s}$ ([2]); that is

$$f(\mathbf{s}) = \sum_{i=1}^{\infty} C(s_i, s_{i+1}), \ \forall \mathbf{s} \in \Pi.$$

We assume that f is uniformly convergent over Π; that is, for any $\varepsilon > 0$ there exists n_ε such that for all trajectories $\mathbf{s}$ the relation $\sum_{i=n}^{\infty} C(s_i, s_{i+1}) < \varepsilon$ holds for all $n \geq n_\varepsilon$. In this case f is continuous on Π. Note that this is not a restrictive assumption; it holds if the cost function $C(s_i, s_{i+1})$ is uniformly bounded and also is discounted, for example, by $(1/2)^i$ (see Assumption 1 and Lemma 2 in [2]). Taking this into account, one can define the total cost in the form

$$f(\mathbf{s}) = \sum_{i=1}^{\infty} r^i C(s_i, s_{i+1}); \tag{2.3}$$

where $r \in (0, 1)$ is a discount factor.

We consider the following optimization problem

$$\text{Minimize} \quad f(\mathbf{s}), \qquad \text{subject to} \qquad \mathbf{s} \in \Omega; \tag{2.4}$$

where $\Omega \subset \Pi$ is a given closed set. We note that Ω is also a compact set. Since f is continuous and Ω is compact, an optimal solution s^* to problem (2.4) exists.

Given $\varepsilon > 0$, the ε-neighborhood of $\mathbf{s}$ in Π is defined by

$$V_\varepsilon(\mathbf{s}) \doteq \{\mathbf{s}' \in \Pi : \ \rho(\mathbf{s}, \mathbf{s}') < \varepsilon\}.$$

Trajectory $\mathbf{s} \in \Omega$ is called an isolated point of Ω if there is $\varepsilon > 0$ such that $(V_\varepsilon(\mathbf{s}) \setminus \{\mathbf{s}\}) \cap \Omega = \emptyset$. Clearly, if $\mathbf{s} \in \Omega$, $\Omega \setminus \{\mathbf{s}\} \neq \emptyset$ and $\mathbf{s}$ is not an isolated point of Ω then there is a sequence $\mathbf{s}^n \in \Omega$ such that $\mathbf{s}^n \neq \mathbf{s}$, $\forall n$, and $\rho(\mathbf{s}^n, \mathbf{s}) \to 0$ as $n \to \infty$.

Trajectory $\mathbf{s} \in \Omega$ is called an interior point of Ω, if $\mathbf{s}$ is not an isolated point of Ω and there is $\varepsilon > 0$ such that $V_\varepsilon(\mathbf{s}) \subset \Omega$.

We note that the set Π is not a linear space. Below we will define the notions of "direction" and then "directional derivative" by using the main idea behind these notions defined in linear spaces. They will be used to derive optimality conditions.

Definition 2.2. We say the trajectories $\mathbf{s}, \mathbf{h} \in \Omega$ have the same direction if there are $n_s, n_h \in \mathbb{N}$ such that $s_{n_s+i} = h_{n_h+i}$ for all $i = 1, 2, \cdots$. We will use the notation $(\mathbf{s})_\infty = (\mathbf{h})_\infty$ in this case.

According to this definition, having the same direction means that these trajectories coincide/join after some finite steps; i.e. n_s and n_t. Moreover, the "larger" numbers n_s and n_t, assuming the sets $\{s_2, \cdots, s_{n_s}\}$ and $\{t_2, \cdots, t_{n_t}\}$ are disjoint, can be interpreted as $\mathbf{s}$ and $\mathbf{h}$ being "far" from each-other.

Definition 2.3. Assume that $\bar{\mathbf{s}}$ is not an isolated point of Ω; that is, $V_\varepsilon(\bar{\mathbf{s}}) \cap \Omega \neq \emptyset$ for all $\varepsilon > 0$. The *upper contingent derivative* of f at $\bar{\mathbf{s}}$ with respect to the direction $\bar{\mathbf{d}} \in \Pi$ is defined as

$$Uf(\bar{\mathbf{s}}, \bar{\mathbf{d}}) = \limsup_{\substack{\mathbf{s} \to \bar{\mathbf{s}} \\ \mathbf{d} \to \bar{\mathbf{d}} \\ \mathbf{s} \neq \bar{\mathbf{s}} \\ (\mathbf{s})_\infty = (\mathbf{d})_\infty}} \frac{f(\mathbf{s}) - f(\bar{\mathbf{s}})}{\rho(\mathbf{s}, \bar{\mathbf{s}})}.$$

The idea behind this definition is that the limit of the fraction on the right hand side is taken over all sequences $\mathbf{s}^k, \mathbf{d}^k$ such that $\mathbf{s}^k \to \bar{\mathbf{s}}$, $\mathbf{d}^k \to \bar{\mathbf{d}}$ and for each k, the trajectory $\mathbf{s}^k$ has the same direction as $\mathbf{d}^k$. Thus, this idea is similar to the definition of the Clarke's directional derivative of g at $\bar{x}$ in direction x in linear spaces given below:

$$g^\circ(\bar{x}, x) = \limsup_{\substack{y \to \bar{x} \\ \xi \downarrow 0}} \frac{f(y + \xi x) - f(y)}{\xi}.$$

Since Π is not a linear space and the operation $\lambda \bar{\mathbf{d}}$, (λ is a given number) is not defined, the super-linearity of $Uf(\bar{\mathbf{s}}, \bar{\mathbf{d}})$ with respect to $\bar{\mathbf{d}}$ can not be considered.

We also note that, according to (2.2), for example if $\rho(\mathbf{s}^k, \bar{\mathbf{s}}) \leq \varepsilon$, then the first $[1 + log_2 \frac{1}{\varepsilon}]$ elements of trajectories $\mathbf{s}^k$ and $\bar{\mathbf{s}}$ coincide. Therefore, the statements $\mathbf{s}^k \to \bar{\mathbf{s}}$, $\mathbf{d}^k \to \bar{\mathbf{d}}$ means that trajectories $\mathbf{s}^k$ and $\mathbf{d}^k$ coincide with $\bar{\mathbf{s}}$ and $\bar{\mathbf{d}}$, respectively, at the some "initial stage", and this "initial stage" grows infinitely when $k \to \infty$. Then, the condition $(\mathbf{s}^k)_\infty = (\mathbf{d}^k)_\infty$ states that trajectories $\mathbf{s}^k$, $\mathbf{d}^k$ join after that "initial stage"; that is, have the same direction in terms of Definition 2.2. Therefore, we can interpret this siltation as trajectories $\mathbf{s}^k$ approaching to $\bar{\mathbf{s}}$ in direction $\bar{\mathbf{d}}$.

Clearly, the upper contingent derivative $Uf(\bar{\mathbf{s}}, \bar{\mathbf{d}})$ is not defined for directions (trajectories) $\bar{\mathbf{d}}$ that are not connected with $\bar{\mathbf{s}}$ in the sense of the above interpretation.

3 Contingent cone

Let $\Omega \subset \Pi$ and $\bar{\mathbf{s}} \in \Omega$ be a non-isolated point of Ω. For the rest of paper we assume that $\Omega \setminus \{\mathbf{s}\} \neq \emptyset$. We introduce the notion of contingent cone $T_\Omega(\bar{\mathbf{s}})$.

Definition 3.1. We say $\mathbf{s} \in \Pi$ is an element of contingent cone $T_\Omega(\bar{\mathbf{s}})$ to the set Ω at $\bar{\mathbf{s}}$ if there exist sequences of trajectories $\mathbf{s}^n \in \Omega, (\mathbf{s}^n \neq \bar{\mathbf{s}}, \forall n)$ and $\mathbf{t}^n \in \Pi$ such that $\mathbf{s}^n \to \bar{\mathbf{s}}$, $\mathbf{t}^n \to \mathbf{s}$ as $n \to \infty$, and $(\mathbf{s}^n)_\infty = (\mathbf{t}^n)_\infty$ for all n :

$$T_\Omega(\bar{\mathbf{s}}) \doteq \{\mathbf{s} \in \Pi : \ \exists \mathbf{s}^n \in \Omega, \mathbf{t}^n \in \Pi; \ \mathbf{s}^n \neq \bar{\mathbf{s}}, \forall n, \ \mathbf{s}^n \to \bar{\mathbf{s}}, \mathbf{t}^n \to \mathbf{s} \text{ and } (\mathbf{s}^n)_\infty = (\mathbf{t}^n)_\infty, \forall n\}.$$

Roughly speaking, the contingent cone $T_\Omega(\bar{\mathbf{s}})$ combines all trajectories in Π having the same direction with some trajectory in Ω being "sufficiently" close to $\bar{\mathbf{s}}$.

Some properties of the contingent cone are presented below.

Lemma 3.2. *Let $\bar{\mathbf{s}}$ be a non-isolated point of Ω. Then $T_\Omega(\bar{\mathbf{s}})$ is not empty and, in particular, $\bar{\mathbf{s}} \in T_\Omega(\bar{\mathbf{s}})$.*

Proof: If $\bar{\mathbf{s}}$ is a non-isolated point of Ω, then there exists a sequence $\mathbf{s}^n \in \Omega$ such that $\mathbf{s}^n \neq \bar{\mathbf{s}}, \forall n$, and $\mathbf{s}^n \to \bar{\mathbf{s}}$ as $n \to \infty$. Then, to prove the relation $\bar{s} \in T_\Omega(\bar{\mathbf{s}})$ it is enough to let $\mathbf{t}^n := \mathbf{s}^n$ in Definition 3.1. $\square$

In the next lemma, we show that the contingent cone in Definition 3.1 is a closed set.

Lemma 3.3. $T_\Omega(\bar{\mathbf{s}})$ *is a closed set.*

Proof: Assume that the sequence $\mathbf{u}^m \in T_\Omega(\bar{\mathbf{s}})$ converges to $\mathbf{u}$. To show $T_\Omega(\bar{\mathbf{s}})$ is closed, it suffices to show that $\mathbf{u} \in T_\Omega(\bar{\mathbf{s}})$; or equivalently, it suffices to show that there exist sequences $\mathbf{s}^n \in \Omega$ ($\mathbf{s}^n \neq \bar{\mathbf{s}}, \forall n$) and $\mathbf{t}^n \in \Pi$ such that the conditions of Definition 3.1 are satisfied; that is

$$\mathbf{s}^n \to \bar{\mathbf{s}}, \quad \mathbf{t}^n \to \mathbf{u} \text{ as } n \to \infty \quad \text{and } (\mathbf{s}^n)_\infty = (\mathbf{t}^n)_\infty, \forall n. \tag{3.1}$$

Take an arbitrary $\varepsilon > 0$. Let $m(\varepsilon)$ such that

$$\rho(\mathbf{u}^{m(\varepsilon)}, \mathbf{u}) < \frac{1}{2}\varepsilon.$$

Given $m(\varepsilon)$, by the definition of $\mathbf{u}^{m(\varepsilon)} \in T_\Omega(\bar{\mathbf{s}})$, there are sequences $\mathbf{s}^{m(\varepsilon),n} \in \Omega$, $\mathbf{t}^{m(\varepsilon),n} \in \Pi$ such that $\mathbf{s}^{m(\varepsilon),n} \neq \bar{\mathbf{s}}, \forall n$, and

$$\mathbf{s}^{m(\varepsilon),n} \to \bar{\mathbf{s}}, \quad \mathbf{t}^{m(\varepsilon),n} \to \mathbf{u}^{m(\varepsilon)} \text{ as } n \to \infty \quad \text{and } (\mathbf{s}^{m(\varepsilon),n})_\infty = (\mathbf{t}^{m(\varepsilon),n})_\infty, \forall n.$$

In other words, there is $n(\varepsilon)$ such that

$$\rho(\mathbf{s}^{m(\varepsilon),n(\varepsilon)}, \bar{\mathbf{s}}) \leq \varepsilon \quad \text{and} \quad \rho(\mathbf{t}^{m(\varepsilon),n(\varepsilon)}, \mathbf{u}^{m(\varepsilon)}) \leq \frac{1}{2}\varepsilon.$$

Then

$$\rho(\mathbf{t}^{m(\varepsilon),n(\varepsilon)}, \mathbf{u}) \leq \rho(\mathbf{t}^{m(\varepsilon),n(\varepsilon)}, \mathbf{u}^{m(\varepsilon)}) + \rho(\mathbf{u}^{m(\varepsilon)}, \mathbf{u}) \leq \varepsilon.$$

This means that

$$\mathbf{s}^{m(\varepsilon),n(\varepsilon)} \to \bar{\mathbf{s}}, \quad \mathbf{t}^{m(\varepsilon),n(\varepsilon)} \to \mathbf{u} \quad \text{as } \varepsilon \to 0.$$

On the other hand $(\mathbf{s}^{m(\varepsilon),n(\varepsilon)})_\infty = (\mathbf{t}^{m(\varepsilon),n(\varepsilon)})_\infty$; that is, the required relation (3.1) is true. $\square$

We note that the sets Ω and $T_\Omega(\bar{\mathbf{s}})$ are generally different. Below we provide an example for which the relation $T_\Omega(\bar{\mathbf{s}}) \nsubseteq \Omega$ holds. A similar example to demonstrate $\Omega \nsubseteq T_\Omega(\bar{\mathbf{s}})$ can be constructed easily.

Example 3.4. Consider the set of nodes $(\delta_1, \delta_2, \delta_3, ...)$, $\delta_i \in \Sigma, \forall i$, and assume that the set

$$\Delta = \{(\delta_{i_1}, \delta_{i_2}, \delta_{i_3}, ...) : \ \{1, 3, 5, \cdots\} \subset \{i_1, i_2, i_3, \cdots\} \text{ and } i_1 < i_2 < \cdots\}$$

belongs to the set of all trajectories Π. Let

$$\bar{\mathbf{s}} = (\delta_1, \delta_3, \delta_5, ...), \quad \mathbf{t} = (\delta_1, \delta_2, \delta_3, \delta_4, ...) \text{ and}$$

$$\Omega = \{\mathbf{s} \in \Delta : \quad (\mathbf{s})_\infty = (\bar{\mathbf{s}})_\infty\}.$$

Clearly $\mathbf{t} \in \Delta \subset \Pi$ and $\mathbf{t} \notin \Omega$ (since each trajectory in Ω contains only odd indices after some finite index). Moreover, $\bar{\mathbf{s}} \in \Omega$ is a non-isolated point of Ω; for example, for the set of trajectories $\mathbf{s}^n \in \Omega$ defined by

$$\mathbf{s}^n = (\delta_1, \delta_3, \delta_5, \cdots, \delta_{2n+1}, \delta_{2n+2}, \delta_{2n+3}, \cdots, \delta_{4n+1}, \delta_{4n+3}, \delta_{4n+5}, ...)$$

we have $\mathbf{s}^n \neq \bar{\mathbf{s}}, \forall n \geq 1$ and $\mathbf{s}^n \to \bar{\mathbf{s}}$ as $n \to \infty$.

We show that $\mathbf{t} \in T_\Omega(\bar{\mathbf{s}})$. Consider the sequence of trajectories $\mathbf{t}^n \in \Delta \subset \Pi$ defined as follows

$$\mathbf{t}^n = (\delta_1, \delta_2, \delta_3, ..., \delta_{2n}, \delta_{2n+1}, \delta_{2n+3}, \delta_{2n+5}, ...).$$

Clearly $\mathbf{t}^n \to \mathbf{t}$ as $n \to \infty$; and moreover, $(\mathbf{t}^n)_\infty = (\mathbf{s}^n)_\infty$ for all n.

Therefore, $\mathbf{t} \in T_\Omega(\bar{\mathbf{s}})$ and $\mathbf{t} \notin \Omega$; that is, $T_\Omega(\bar{\mathbf{s}}) \not\subseteq \Omega$. $\square$

In the following, we investigate the contingent cone of intersection and union of sets.

Lemma 3.5. *Let Ω_1 and Ω_2 be subsets of Π and $\bar{\mathbf{s}} \in \Omega_1 \cap \Omega_2$. Then,*

(i) $T_{\Omega_1 \cap \Omega_2}(\bar{\mathbf{s}}) \subset T_{\Omega_1}(\bar{\mathbf{s}}) \cap T_{\Omega_2}(\bar{\mathbf{s}})$.

(ii) $T_{\Omega_1 \cup \Omega_2}(\bar{\mathbf{s}}) = T_{\Omega_1}(\bar{\mathbf{s}}) \cup T_{\Omega_2}(\bar{\mathbf{s}})$.

Proof: The proofs of $T_{\Omega_1 \cap \Omega_2}(\bar{\mathbf{s}}) \subset T_{\Omega_1}(\bar{\mathbf{s}}) \cap T_{\Omega_2}(\bar{\mathbf{s}})$ and $T_{\Omega_1}(\bar{\mathbf{s}}) \cup T_{\Omega_2}(\bar{\mathbf{s}}) \subset T_{\Omega_1 \cup \Omega_2}(\bar{\mathbf{s}})$ directly follow from the definition of contingent cone. Thus, it is enough to show that $T_{\Omega_1 \cup \Omega_2}(\bar{\mathbf{s}}) \subset T_{\Omega_1}(\bar{\mathbf{s}}) \cup T_{\Omega_2}(\bar{\mathbf{s}})$.

Let $\mathbf{t} \in T_{\Omega_1 \cup \Omega_2}(\bar{\mathbf{s}})$. Then, there are sequences $\mathbf{s}^n \in \Omega_1 \cup \Omega_2$ and $\{\mathbf{t}^n\}_{n \in \mathbb{N}}$ such that

$$\rho(\mathbf{s}^n, \bar{\mathbf{s}}) \to 0, \rho(\mathbf{t}^n, \mathbf{t}) \to 0, \text{ and } (\mathbf{s}^n)_\infty = (\mathbf{t}^n)_\infty, \forall n. \tag{3.2}$$

As $\mathbf{s}^n \in \Omega_1 \cup \Omega_2$, for some $j \in \{1, 2\}$ the realtion $\mathbf{s}^n \in \Omega_j$ holds. Then, from (3.2) we conclude that $\mathbf{t} \in T_{\Omega_j}(\bar{\mathbf{s}})$ and hence $\mathbf{t} \in T_{\Omega_1}(\bar{\mathbf{s}}) \cup T_{\Omega_2}(\bar{\mathbf{s}})$. $\square$

In the next example, we show that the inverse of the inclusion **(i)** in this lemma; that is, the relation

$$T_{\Omega_1 \cap \Omega_2}(\bar{\mathbf{s}}) \supset T_{\Omega_1}(\bar{\mathbf{s}}) \cap T_{\Omega_2}(\bar{\mathbf{s}})$$

may not be true.

Example 3.6. Consider two trajectories $\mathbf{t} = (\bar{s}_1, t_2, t_3, ...)$ and $\bar{\mathbf{s}} = (\bar{s}_1, \bar{s}_2, \bar{s}_3, ...)$ assuming that $t_i \neq \bar{s}_j$ for all $i, j \geq 2$. Define the sets Ω_1 and Ω_2 as follows

$$\Omega_1 = \{\bar{\mathbf{s}}\} \cup \Omega_{1,2} \cup \Omega_1^0, \quad \Omega_2 = \{\bar{\mathbf{s}}\} \cup \Omega_{1,2} \cup \Omega_2^0;$$

where

$$\Omega_1^0 = \{\mathbf{s}^n = (\bar{s}_1, \bar{s}_2, ..., \bar{s}_{2n}, \bar{s}_{2n+2}, \bar{s}_{2n+4}, \bar{s}_{2n+6}, ...), \quad n = 1, 2, \cdots\},$$

$$\Omega_2^0 = \{\mathbf{u}^n = (\bar{s}_1, \bar{s}_2, ..., \bar{s}_{2n}, \bar{s}_{2n+1}, \bar{s}_{2n+3}, \bar{s}_{2n+5}, ...), \quad n = 1, 2, \cdots\},$$

$$\Omega_{1,2} = \{\mathbf{u}^n = (\bar{s}_1, \bar{s}_2, ..., \bar{s}_{2n}, \xi_{2n+1}, \xi_{2n+2}, \xi_{2n+3}, ...), \quad n = 1, 2, \cdots\};$$

and $\xi_i \neq \bar{s}_j$, for $\xi_i \neq t_j$, all $i, j \geq 2$.

Now, let the set all trajectories Π is given by

$$\Pi = \{\mathbf{t}\} \cup \Omega_1 \cup \Omega_2 \cup T_1 \cup T_2$$

where

$$T_1 = \{\mathbf{t}^n = (t_1, t_2, ..., t_{2n}, \bar{s}_{2n+2}, \bar{s}_{2n+4}, \bar{s}_{2n+6}, ...), \quad n = 1, 2, \cdots\};$$

$$T_2 = \{\mathbf{v}^n = (t_1, t_2, ..., t_{2n}, \bar{s}_{2n+1}, \bar{s}_{2n+3}, \bar{s}_{2n+5}, ...), \quad n = 1, 2, \cdots\}.$$

First we show that $\mathbf{t} \in T_{\Omega_1}(\bar{\mathbf{s}}) \cap T_{\Omega_2}(\bar{\mathbf{s}})$.

Consider the sequences $\mathbf{s}^n \in \Omega_1$ and $\mathbf{t}^n \in T_1$. Clearly

$$\rho(\mathbf{s}^n, \bar{\mathbf{s}}) \to 0, \ \rho(\mathbf{t}^n, \mathbf{t}) \to 0, \ \text{and } (\mathbf{s}^n)_\infty = (\mathbf{t}^n)_\infty, \ \forall n.$$

This means that $\mathbf{t} \in T_{\Omega_1}(\bar{\mathbf{s}})$.

In a similar way, for sequences $\mathbf{u}^n \in \Omega_2$ and $\mathbf{v}^n \in T_2$ we have

$$\rho(\mathbf{u}^n, \bar{\mathbf{s}}) \to 0, \ \rho(\mathbf{v}^n, \mathbf{t}) \to 0, \ \text{and } (\mathbf{u}^n)_\infty = (\mathbf{v}^n)_\infty, \ \forall n;$$

that leads to $\mathbf{t} \in T_{\Omega_2}(\bar{\mathbf{s}})$. Therefore, $\mathbf{t} \in T_{\Omega_1}(\bar{\mathbf{s}}) \cap T_{\Omega_2}(\bar{\mathbf{s}})$.

Now we show that $\mathbf{t} \notin T_{\Omega_1 \cap \Omega_2}(\bar{\mathbf{s}})$. By contradiction let $\mathbf{t} \in T_{\Omega_1 \cap \Omega_2}(\bar{\mathbf{s}})$; that is, there are sequences $\mathbf{u}^n \in \Omega_1 \cap \Omega_2$, $\mathbf{u}^n \neq \bar{\mathbf{s}}, \forall n$, and $\mathbf{d}^n \in \Pi$ such that

$$\rho(\mathbf{u}^n, \bar{\mathbf{s}}) \to 0, \ \rho(\mathbf{d}^n, \mathbf{t}) \to 0, \ \text{and } (\mathbf{u}^n)_\infty = (\mathbf{d}^n)_\infty, \ \forall n. \tag{3.3}$$

As $\Omega_1 \cap \Omega_2 = \{\bar{\mathbf{s}}\} \cup \Omega_{1,2}$, it is not difficult to observe that the relation $\mathbf{u}^n \in \Omega_1 \cap \Omega_2$, $\mathbf{u}^n \neq \bar{\mathbf{s}}$, implies for every n the following holds

$$(\mathbf{u}^n)_\infty \neq (s)_\infty, \quad \forall s \in \{\mathbf{t}\} \cup \Omega_1^0 \cup \Omega_2^0 \cup T_1 \cup T_2.$$

On the other hand, according to the definition of sets Ω_i and T_i, $i = 1, 2$, the convergence $\mathbf{d}^n \to \mathbf{t}$ implies that given any n, one of the following holds: $\mathbf{d}^n = \mathbf{t}$, or $\mathbf{d}^n \in T_1$, or $\mathbf{d}^n \in T_2$. In all of these three cases we have

$$(\mathbf{d}^n)_\infty \neq (\mathbf{u}^n)_\infty.$$

This contradicts (3.3). Thus, $\mathbf{t} \notin T_{\Omega_1 \cap \Omega_2}(\bar{\mathbf{s}})$. □

4 Optimality conditions

Consider the optimization problem (2.4) stated in Section 2. In this section our aim is to investigate optimality conditions for this problem in terms of the contingent cone and the upper contingent derivative.

Assumption (L): The total cost f is locally Lipschitz; that is, for each $\bar{\mathbf{s}} \in \Pi$, there exist a $\delta-$neighborhood $V_\delta(\bar{\mathbf{s}})$ with $\delta > 0$ and a number $L_{\delta,\bar{\mathbf{s}}} < \infty$ such that

$$|f(\mathbf{s}) - f(\mathbf{t})| \leq L_{\delta,\bar{\mathbf{s}}} \, \rho(\mathbf{s}, \mathbf{t}), \ \forall \mathbf{s}, \mathbf{t} \in V_\delta(\bar{\mathbf{s}}). \tag{4.1}$$

The following two propositions provide some sufficient conditions under which Assumption (L) holds.

Proposition 4.1. *Assume that cost function C is bounded and the total cost f is defined as in (2.3). If $r \leq \frac{1}{2}$ then f is locally Lipschitz at each $\bar{\mathbf{s}} \in \Pi$.*

Proof: Given $\bar{\mathbf{s}} \in \Pi$, we take any $\delta > 0$ and trajectories $\mathbf{s}, \mathbf{t} \in V_\delta(\bar{\mathbf{s}})$. Clearly

$$\rho(\mathbf{s}, \mathbf{t}) \leq 2\delta.$$

According to (2.2) at least the first $[log_2 \frac{1}{2\delta} - 1]$ elements of trajectories $\mathbf{s} = (s_1, s_2, \cdots)$ and $\mathbf{t} = (t_1, t_2, \cdots)$ coincide. Let $N_\delta \geq [log_2 \frac{1}{2\delta} - 1]$ be the first index for which $s_{N_\delta} \neq t_{N_\delta}$ and $s_i = t_i$ for all $i = 1, \cdots, N_\delta - 1$. Then from (2.1) we have

$$\rho(\mathbf{s}, \mathbf{t}) = \sum_{i=1}^{\infty} \phi_i(\mathbf{s}, \mathbf{t}) 2^{-i} \geq 2^{-N_\delta}. \tag{4.2}$$

On the other hand

$$|f(\mathbf{s})-f(\mathbf{t})| \leq \sum_{i=1}^{\infty} r^i |C(s_i,s_{i+1})-C(t_i,t_{i+1})| = \sum_{i=N_\delta-2}^{\infty} r^i |C(s_i,s_{i+1})-C(t_i,t_{i+1})|.$$

Since cost function C is bounded, there is $M<\infty$ such that $|C(s_i,s_{i+1})| \leq M$ for all i and $\mathbf{s}$. Then

$$|f(\mathbf{s})-f(\mathbf{t})| \leq \sum_{i=N_\delta-2}^{\infty} 2\, r^i M \leq 2Mr^{N_\delta-2}\frac{1}{1-r}.$$

Taking into account $r \leq \frac{1}{2}$ we have $\frac{1}{1-r} \leq 2$ and $r^{N_\delta-2} \leq 2^{-N_\delta+2}$ and therefore

$$|f(\mathbf{s})-f(\mathbf{t})| \leq 2M\,2^{-N_\delta+1}.$$

Thus, denoting $L_{\delta,\bar{\mathbf{s}}} = 4M$ from (4.2) we obtain (4.1). □

Proposition 4.2. *Let $\bar{\mathbf{s}} \in \Pi$ be given and the total cost f be defined as in (2.3). Assume that, there exist a $\delta-$neighborhood $V_\delta(\bar{\mathbf{s}})$ and a number $M<\infty$ such that the cost function C satisfyies the following condition*

$$|C(s_i,s_{i+1})-C(t_i,t_{i+1})| \leq M\,\rho(\mathbf{s},\mathbf{t}),\ \forall \mathbf{s},\mathbf{t} \in V_\delta(\bar{\mathbf{s}}) \text{ and } \forall i=1,2,3,\dots .$$

Then f is locally Lipschitz at $\bar{\mathbf{s}}$.

Proof: Take any $\mathbf{s},\mathbf{t} \in V_\delta(\bar{\mathbf{s}})$. We have

$$|f(\mathbf{s})-f(\mathbf{t})| \leq \sum_{i=1}^{\infty} r^i |C(s_i,s_{i+1})-C(t_i,t_{i+1})|$$

$$\leq \sum_{i=1}^{\infty} r^i\, M\,\rho(\mathbf{s},\mathbf{t}) = \frac{Mr}{1-r}\,\rho(\mathbf{s},\mathbf{t}),\ \forall \mathbf{s} \in V_\delta(\mathbf{t}).$$

Thus, (4.1) holds for $L_{\delta,\bar{\mathbf{s}}} = \frac{Mr}{1-r}$. □

The following results is about the existence of the upper contingent derivative.

Lemma 4.3. *Assume that $\bar{\mathbf{s}}$ is a non-isolated point of Ω and Assumption (L) holds. Then the upper contingent derivative $Uf(\bar{\mathbf{s}},\mathbf{d})$ exists for each direction $\mathbf{d} \in T_\Omega(\bar{\mathbf{s}}) \setminus \{\bar{\mathbf{s}}\}$.*

Proof: Take an arbitrary $\bar{\mathbf{d}} \in T_\Omega(\bar{\mathbf{s}}) \setminus \{\bar{\mathbf{s}}\}$. By Assumption (L) for $\bar{\mathbf{s}} \in \Omega$, there exists a $\delta-$neighborhood $V_\delta(\bar{\mathbf{s}})$ with $\delta > 0$ such that (4.1) holds. Then

$$\frac{|f(\mathbf{s}) - f(\bar{\mathbf{s}})|}{\rho(\mathbf{s}, \bar{\mathbf{s}})} \le L_{\delta,\bar{\mathbf{s}}}, \ \forall \mathbf{s} \in V_\delta(\bar{\mathbf{s}}) \setminus \{\bar{\mathbf{s}}\}. \tag{4.3}$$

Since $\bar{\mathbf{d}} \in T_\Omega(\bar{\mathbf{s}})$, then, it follows from (4.3) that

$$-L_{\delta,\bar{\mathbf{s}}} \le \sup_{\substack{\rho(\mathbf{s},\bar{\mathbf{s}})<\varepsilon \\ \rho(\mathbf{d},\bar{\mathbf{d}})<\varepsilon \\ \mathbf{s}\neq\bar{\mathbf{s}} \\ (\mathbf{s})_\infty=(\mathbf{d})_\infty}} \frac{f(\mathbf{s}) - f(\bar{\mathbf{s}})}{\rho(\mathbf{s}, \bar{\mathbf{s}})} \le L_{\delta,\bar{\mathbf{s}}}, \quad \forall \epsilon < \delta. \tag{4.4}$$

Hence

$$Uf(\bar{\mathbf{s}}, \bar{\mathbf{d}}) = \limsup_{\substack{\mathbf{s}\to\bar{\mathbf{s}} \\ \mathbf{d}\to\bar{\mathbf{d}} \\ \mathbf{s}\neq\bar{\mathbf{s}} \\ (\mathbf{s})_\infty=(\mathbf{d})_\infty}} \frac{f(\mathbf{s}) - f(\bar{\mathbf{s}})}{\rho(\mathbf{s}, \bar{\mathbf{s}})}$$

exists, which means, the upper contingent derivative exists for any non-isolated point $\bar{\mathbf{s}} \in \Omega$ and any direction $\bar{\mathbf{d}} \in T_\Omega(\bar{\mathbf{s}}) \setminus \{\bar{\mathbf{s}}\}$. □

Remark: According to Lemma 4.3, as $Uf(\bar{\mathbf{s}}, \mathbf{d})$ exists if $\bar{\mathbf{d}}$ is in contingent cone, we call it the upper contingent derivative.

We will call d a descent direction of f at $\bar{\mathbf{s}}$ if the upper contingent derivative is negative: $Uf(\bar{\mathbf{s}}, \mathbf{d}) < 0$. The main result of this section is a necessary condition of optimality for a local minimizer of f that are presented in the next theorem and corollary.

Theorem 4.4. *Assume that s^* is a non-isolated point of Ω and Assumption (L) holds at $\mathbf{s}^*$. If $Uf(\mathbf{s}^*, \mathbf{d}) < 0$ for some $\mathbf{d} \in T_\Omega(\mathbf{s}^*) \setminus \{\bar{\mathbf{s}}\}$ then, $\mathbf{s}^*$ is not a local minimizer of the problem (2.4); that is, for every $\varepsilon > 0$, there exist $\mathbf{s}^\varepsilon \in \Omega$ and $\mathbf{d}^\varepsilon \in \Pi$ such that*

$$\rho(\mathbf{s}^\varepsilon, \mathbf{s}^*) < \varepsilon, \ \rho(\mathbf{d}^\varepsilon, \mathbf{d}) < \varepsilon, \ \mathbf{s}^\varepsilon \neq \mathbf{s}^*, \ (\mathbf{s}^\varepsilon)_\infty = (\mathbf{d}^\varepsilon)_\infty \ \text{ and } \ f(\mathbf{s}^\varepsilon) < f(\mathbf{s}^*). \tag{4.5}$$

Proof: By the assumption, there exists $\bar{\mathbf{d}} \in T_\Omega(\mathbf{s}^*) \setminus \{\mathbf{s}^*\}$ such that $Uf(\mathbf{s}^*, \bar{\mathbf{d}}) < 0$. Since $\bar{\mathbf{d}} \in T_\Omega(\bar{\mathbf{s}})$, by the definition of contingent cone, there exist sequences of trajectories $\mathbf{s}^n \in \Omega$ and $\mathbf{t}^n \in \Pi$ such that

$$\mathbf{s}^n \to \mathbf{s}^*, \mathbf{t}^n \to \bar{\mathbf{d}} \text{ as } n \to \infty; \text{ and } \mathbf{s}^n \neq \mathbf{s}^*, \ (\mathbf{s}^n)_\infty = (\mathbf{d}^n)_\infty, \forall n \ge 1. \tag{4.6}$$

By the definition of upper contingent derivative we have

$$\limsup_{n\to\infty} \frac{f(\mathbf{s}^n) - f(\mathbf{s}^*)}{\rho(\mathbf{s}^n, \mathbf{s}^*)} \leq Uf(\mathbf{s}^*, \bar{\mathbf{d}}) < 0.$$

Then, given $\epsilon > 0$, there exists a sufficiently large number n_ϵ such that the inequalities $\rho(\mathbf{s}^{n_\epsilon}, \mathbf{s}^*) < \epsilon$, $\rho(\mathbf{t}^{n_\epsilon}, \bar{\mathbf{d}}) < \epsilon$ and $f(\mathbf{s}^{n_\epsilon}) - f(\mathbf{s}^*) < 0$ hold. Therefore, (4.6) yields (4.5). □

From Theorem 4.4 we obtain the following necessary condition of local optimality.

Corollary 4.5. (Necessary condition of optimality) *Assume that s^* is a non-isolated point of Ω and Assumption (L) holds at $\mathbf{s}^*$. If $\mathbf{s}^*$ is a local minimizer of the problem (2.4) then,*

$$Uf(\mathbf{s}^*, \mathbf{d}) \geq 0, \quad \forall \mathbf{d} \in T_\Omega(\mathbf{s}^*) \setminus \{\mathbf{s}^*\}. \tag{4.7}$$

In the following, we present an example that illustrates Theorem 4.4 and Corollary 4.5.

Example 4.6. Let

$$\mathbf{s}^* = (s_1, s_2, s_3, s_4 ....), \ \bar{\mathbf{d}} = (s_1, \delta_2, \delta_3, \delta_4, ...), \text{ and}$$

$$\mathbf{s}^n = (s_1, s_2, ..., s_n, \delta_{n+1}, \delta_{n+2}, \delta_{n+3}, ...) : \ n = 1, 2, \cdots ;$$

where $s_i \neq \delta_i$ for all $i \geq 2$.

The set Ω and the set of all trajectories Π are given by $\Omega = \{\mathbf{s}^*\} \cup \{\mathbf{s}^n : \ n = 1, 2, \cdots\}$ and $\Pi = \{\bar{\mathbf{d}}\} \cup \Omega$.

By the definition of Ω for any given $\mathbf{s}^n \in \Omega$ we have

$$\rho(\mathbf{s}^n, \mathbf{s}^*) = \frac{1}{2^{n+1}} + \frac{1}{2^{n+2}} + \cdots = \frac{1}{2^n}. \tag{4.8}$$

Since $\rho(\mathbf{s}^n, \mathbf{s}^*) \to 0$, the trajectory $\mathbf{s}^*$ is not an isolated point of Ω.

The cost function C is defined as follows: for all $i \in \mathbb{N}$,

$$C(\delta_i, s_{i+1}) = \frac{1 - \xi(i)}{2^i}, \text{ and } C(s_i, s_{i+1}) = C(\delta_i, \delta_{i+1}) = \frac{1}{2^i};$$

where $\xi(i)$ satisfies

$$|\xi(i)| \leq M < \infty, \forall i \geq 1, \text{ and } \lim_{i\to\infty} \xi(i) = \xi^*. \tag{4.9}$$

Therefore, we have

$$f(\mathbf{s}^*) = \sum_{i=1}^{\infty} \frac{1}{2^i} = 1 \text{ and } f(\mathbf{s}^n) = 1 - \frac{\xi(n)}{2^n}, \forall n \geq 1. \tag{4.10}$$

Now, we show that Assumption (L) holds at $\mathbf{s}^*$. Let $\mathbf{s}$ and $\mathbf{t}$ in Ω be arbitrary. Then, $\mathbf{s} = \mathbf{s}^n$ and $\mathbf{t} = \mathbf{t}^m$ for some $n, m \in \mathbb{N}$. Assume that $n < m$. We have

$$|f(\mathbf{s}) - f(\mathbf{t})| = \left| -\frac{\xi(n)}{2^n} + \frac{\xi(m)}{2^m} \right| \leq M \left(\frac{1}{2^n} + \frac{1}{2^m} \right).$$

By the fact that

$$\rho(\mathbf{s}, \mathbf{t}) \geq 1/2^n + 1/2^{n+1} + ... + 1/2^m \geq 1/2^n + 1/2^m,$$

we obtain

$$|f(\mathbf{s}) - f(\mathbf{t})| \leq M\, \rho(\mathbf{s}, \mathbf{t}).$$

This implies that Assumption (L) holds at $\mathbf{s}^*$.

Consider arbitrary sequences of trajectories $\mathbf{s}^{n_k} \in \Omega$ and $\mathbf{d}^{n_k} \to \bar{\mathbf{d}}$. First we note that $\mathbf{d}^{n_k} \to \bar{\mathbf{d}}$ implies $\mathbf{d}^{n_k} = \bar{\mathbf{d}}, \forall n_k$. On the other hand, by the definition of Ω it is not not difficult to observe that $\mathbf{s}^{n_k} \to \bar{\mathbf{s}}$ as $n_k \to \infty$. Moreover, $(\mathbf{s}^{n_k})_\infty = (\mathbf{d}^{n_k})_\infty, \forall n_k$. Thus, $\bar{\mathbf{d}} \in T_\Omega(\mathbf{s}^*) \setminus \{\bar{\mathbf{s}}\}$.

By Lemma (4.3) the upper contingent derivative $Uf(\mathbf{s}^*, \bar{\mathbf{d}})$ exists for the direction $\bar{\mathbf{d}} \in T_\Omega(\mathbf{s}^*) \setminus \{\mathbf{s}^*\}$. We have

$$Uf(\mathbf{s}^*, \bar{\mathbf{d}}) = \limsup_{\substack{\mathbf{s} \to \mathbf{s}^* \\ \mathbf{d} \to \bar{\mathbf{d}} \\ \mathbf{s} \neq \mathbf{s}^* \\ (\mathbf{s})_\infty = (\mathbf{d})_\infty}} \frac{f(\mathbf{s}) - f(\mathbf{s}^*)}{\rho(\mathbf{s}, \mathbf{s}^*)} = \limsup_{n_k \to \infty} \frac{f(\mathbf{s}^{n_k}) - f(\mathbf{s}^*)}{\rho(\mathbf{s}^{n_k}, \mathbf{s}^*)}$$

and therefore from (4.8), (4.9) and (4.10) it follows that

$$Uf(\mathbf{s}^*, \bar{\mathbf{d}}) = \limsup_{n_k \to \infty} \frac{1 - \frac{\xi(n_k)}{2^{n_k}} - 1}{\frac{1}{2^{n_k}}} = -\xi^*.$$

If $\mathbf{s}^*$ is a local minimizer in the problem (2.4) then, for all sufficiently large n_k the inequality $f(\mathbf{s}^{n_k}) \geq f(\mathbf{s}^*)$ holds and according to (4.10) we have $\xi(n_k) \leq 0$. Thus in this case $\xi^* \leq 0$ or $Uf(\mathbf{s}^*, \bar{\mathbf{d}}) \geq 0$.

Inversely, if $Uf(\mathbf{s}^*, \bar{\mathbf{d}}) < 0$ for some $\bar{\mathbf{d}}$, then $\xi^* > 0$ and the inequality $\xi(n_k) > 0$ holds for sufficiently large numbers n_k. Therefore in this case, there is n_k such that $f(\mathbf{s}^{n_k}) < f(\mathbf{s}^*)$; that is, $\mathbf{s}^*$ is not not a local minimizer of f. $\square$

References

[1] J. Bean and R. Smith. Conditions for the existence of planning horizons. *Math. Oper. Res.*, (9):391–401, 1984.

[2] J. C. Bean and R. L. Smith. Conditions for the discovery of solution horizons. *Math. Program.*, 59:215–229, 1993.

[3] D. P. Bertsekas. *Dynamic Programming and Stochastic Control*. Academic Press, New-York, 1976.

[4] D. Blackweil. Discrete dynamic programming. *Ann. Math. Statist.*, 33:719–726, 1962.

[5] G. Bouligand. Sur les surfaces dépourvues de points hyperlimites. *Ann. Soc. Polon. Math.*, 9:32–41, 1930.

[6] D. A. Carlson, A. Haurie, and A. Leizarowitz. *Infinite Horizon Optimal Control: Deterministic and Stochastic Systems*. 2nd ed. Springer-Verlag, New York, 1991.

[7] F. H. Clarke. *Optimization and nonsmooth analysis*. Wiley, 1983.

[8] R. Dorfman, P. Samuelson, and R. Solow. *Linear programming and economic analysis*. New York: McGraw-Hill, 1958.

[9] A. Y. Dubovitskii and A. Milyiutin. Extremum problems in the presence of constraints, [in Russian]. *Dokl. Akad. Nauk. SSSR*, 149:759–762, 1963.

[10] A. Y. Dubovitskii and A. Milyiutin. Extremum problems in the presence of restrictions, [in Russian]. *USSR Comput. Math. Phys.*, 5:1–80, 1965.

[11] V. Gaitsgory and S. Rossomakhine. Linear programming approach to deterministic long run average problems of optimal control. *SIAM J. Control Optim.*, 44, 2006.

[12] D. Gale. On optimal development in a multi-sector economy. *The Review of Economic Studies*, 34:1–18, 1976.

[13] H. Halkin. Necessary conditions for optimal control problems with infinite horizons. *Econometrica*, 42:267–272, 1974.

[14] W. Hopp, J. Bean, and R. Smith. A new optimality criterion for non-homegeneous markov decesion process. *Oper. Res.*, (35):875–883, 1987.

[15] J. B. Lasserre. Decision horizon, overtaking and 1-optimality criteria in optimal control. *Advances in Optimization. Lecture Notes in Mathematics, no. 302, Eiselt, H. A. and G. Pederzoli, eds. Springer Verlag, New York*, pages 247–261, 1988.

[16] A. Leizarowitz. Infinite horizon autonomous systems with unbounded cost. *Appl. Math. Optim.*, 13:19–43, 1985.

[17] V. Makarov and A. Rubinov. *Mathematical theory of economic dynamics and equilibria.* Springer-Verlag, New York, 1977.

[18] L. W. McKenzie. Turnpike theory. *Econometrica*, 44:841–866, 1976.

[19] S. Rayan. *Degenercy in discrete infinite horizon optimization.* PhD thesis, Department of industrial and operations engineering, The University of Michigan, Ann Arbor, 1988.

[20] S. M. Rayan, J. C. Bean, and R. L. Smith. A tie-breaking rule for discrete infinite horizon optimization. *Oper. Res.*, 40(1):117–126, 1992.

[21] S. M. Rayan, J. C. Bean, and R. L. Smith. A tie-breaking rule for discrete infinite horizon optimization. *Oper. Res.*, 40:117–126, 1992.

[22] I. Schochetman and R. Smith. Existence and discovery of average optimal solutions in deterministic infinite horizon optimization. *Math. Oper. Res.*, pages 416–432, 1998.

[23] I. Schochetman and R. Smith. Optimality criteria for deterministic discrete-time infinite horizon optimization. *Int. J. Math. Math. Sci.*, pages 57–80, 2004.

[24] I. Schochetman and R. Smith. Existence of efficient solutions in infinite horizon optimization under continuous and discrete controls. *Oper. Res. Lett.*, (33):97–104, 2005.

[25] I. E. Schochetman and R. L. Smith. Infinite horizon optimization. *Math. Oper. Res.*, 14(3):559–574, 1989.

[26] I. E. Schochetman and R. L. Smith. Existence and discovery of average optimal solutions in deterministic infinite horizon optimization. *Math. Oper. Res.*, pages 416–432, 1998.

[27] D. Skilton. Imbedding posets in the integers. *Order*, 1(3):229–233, 1985.

[28] A. F. Veinott. On finding optimal policies in discrete dynamic programming with no discounting. *Ann. Math. Statist.*, 37:1284–1294, 1966.

[29] A. O. Wachs, I. E. Schochetman, and R. L. Smith. Average optimality in non-homogeneous infinite horizon markov decision processes. *Math. Oper. Res.*, 36:147–164, 2011.

[30] C. C. Weizsacker. Existence of optimal programs of accumulation for an infinite horizon. *The Review of Economic Studies*, 32:84–105, 1965.

[31] A. Zaslavski. *Turnpike properties in the calculus of variations and optimal control*. Springer, 2006.

Chapter 6

Conclusions and Future Work

Conclusion

In this thesis, we investigate optimality conditions for a broad class of optimization problems involving non-convex non-smooth functions. The main techniques used throughout the thesis deal with the notions of weak subdifferentials, augmented normal cones and superdifferentials.

To start with, several necessary and sufficient optimality conditions involving the Hadamard directional derivative are presented (see paper 1 in section 1.5). Then, a new notion of local sigma supporting cone, based on the definition of augmented normal cone, is introduced. By using this cone two new concepts, namely, the conic gap and the maximal conic gap are introduced. These concepts are applied to investigate if a set has a conic gap at some boundary points and moreover to measure how big such a gap is. The main focus in introducing sigma supporting cone is the optimality conditions for local optimal solutions. These optimality conditions for a special class of problems are presented by using the local supporting functions together with the weak subdifferentials (papers 2 and 3 in section 1.5). Afterwards, we investigate global optimality by introducing the global supporting cone, similar to the local one. Then, all these new concepts and optimality conditions are considered in reflexive Banach spaces (see paper 4,5 and 6 in section 1.5).

Two important applications considered in the thesis deal with the optimality conditions for the difference of topical functions, as well as for the cost function over an infinite horizon. Optimization problems where the objective function is a topical function have been frequently investigated recently. In this thesis, we establish necessary and sufficient optimality conditions where the objective function is defined as a difference of two topical functions (see paper 7 in section 5.1). Infinite horizon optimization is another important class of applications considered in this thesis. The study of stability of optimal trajectories is one of the main concerns of this study. We establish stability of a sequence of minimizing trajectories in terms of the classical and the ideal convergence (see paper 8 in section 5.2). Finally, by introducing a new concept of the contingent derivative (paper 9, section 5.3), optimality conditions for this class of optimization problems are derived.

Future work

There are still many interesting problems related to optimality conditions considered in this thesis that could be investigated. For example, it would be interesting to replace the difference of topical functions in paper 7 in chapter 5 with quotient of two topical functions and derive optimality conditions for this type of functions by applying superdifferentials.

Another interesting problem is related to the optimality conditions in the problem of cost minimization over infinite horizon optimization. These conditions could be derived by introducing some generalization of subdifferentials of cost function.

Bibliography

[1] J. P. Aubin. Gradient génŕalisś de clarke. *Ann. Sci. Math.*, 2:197–253, 1978.

[2] A. Y. Azimov and R. N. Gasimov. On conjugacy, weak subdifferentials and duality with zero gap in nonconvex optimization. *Int. J. Appl. Math*, 1:171–192, 1999.

[3] A. Y. Azimov and R. N. Gasimov. Stability and duality of nonconvex problems via augmented lagrangian. *Cybernetics and Systems Analysis*, 38:412–421, 2002.

[4] F. Baccelli and J. Mairesse. Ergodic theorems for stochastic operators and discrete event networks. *Idempotency (Bristol, 1994)*, 11:171–208, 1998.

[5] V. Barbu and T. Precupanu. *Convexity and Optimization in Banach Spaces*. Springer Monographs in Mathematics. Springer Netherlands, 2012.

[6] H. Barsam and H. Mohebi. *Characterizations of upward and downward sets in semimodules by using topical functions, Numerical Functional Analysis and Optimization.* to appear, 2016.

[7] M. S. Bazaraa, J. J. Goode, and M. Z. Nashed. On the cones of tangents with applications to mathematical programming. *Journal of Optimization Theory and Applications*, 13:389–426, 1974.

[8] J. C. Bean, J. R. Birge, and R. L. Smith. *Aggregation in dynamic programming*, volume 35. INFORMS, 1987.

[9] J. C. Bean and R. L. Smith. Conditions for the existence of planning horizons. *Mathematics of Operations Research*, 9:391–401, 1984.

[10] J. C. Bean and R. L. Smith. *Conditions for the Discovery of Solution Horizons*, volume 59. Springer-Verlag New York, Inc., 1993.

[11] A. Ben-Tal and J. Zowe. Directional derivatives in nonsmooth optimization. *J. Optim. Theory App*, 47:483–490, 1985.

[12] D. P. Bertsekas. *Dynamic Programming and Stochastic Control.* Academic Press, New-York, 1976.

[13] D. Blackweil. Discrete dynamic programming. *Ann. Math. Statist.*, 33:719–726, 1962.

[14] J. M. Borwein and A. S. Lewis. *Convex Analysis And Nonlinear Optimization: Theory And Examples, CMS Books in Mathematics.* Springer, 2006.

[15] G. Bouligand. Sur les surfaces dépourvues de points hyperlimites. *Ann. Soc. Polon. Math*, 9:32–41, 1930.

[16] G. Bouligand. *Introduction á la géométrie infinitésimale directe.* Gauthier-Villars, 1932.

[17] G. Bouligand. *Sur la semi-continuité dínclusions et quelques sujets conn exes*, volume 31. Enseignement Mathématique, 1932.

[18] A. Brondsted and R. T. Rockafellar. On the subdifferentiability of convex functions. *Bull. Amer. Math. Soc.*, 16:605–611, 1965.

[19] D. Carlson, A. B. Haurie, and A. Leizarowitz. *Infinite horizon optimal control: deterministic and stochastic systems.* Springer Science & Business Media, 2012.

[20] G. Chen, X. Huang, and X. Yang. Vector optimization: set-valued and variational analysis. 541, 2006.

[21] F. H. Clarke. *Necessary conditions for nonsmooth problems in optimal control and the calculus of variations.* Department of Mathematics, University of Washington, Seattle, 1973.

[22] F. H. Clarke. Generalized gradients and applications. *Trans. Amer. Math. Soc*, 205:247–262, 1975.

[23] F. H. Clarke. The maximum principle under minimal hypotheses. *Control Optimization*, 14:1078–1091, 1976.

[24] F. H. Clarke. A new approach to lagrange multipliers. *Math of Operations Research*, 1:165–174, 1976.

[25] F. H. Clarke. Nonsmooth analysis and optimization. *Proceedings International Congress of Mathematicians, Helsinki*, 1978.

[26] F. H. Clarke. Generalized gradients of lipschitz functionals. *Adv. Math*, 40:52–67, 1981.

[27] F. H. Clarke. *Optimization and nonsmooth analysis.* Wiley, 1983.

[28] F. H. Clarke and J. P. Aubin. Monotone invariant solutions to differential inclusions. *J. London Math. Soc*, 16:357–366, 1977.

[29] R. A. Cuninghame-Green. Minimax algebra. *Lecture Notes in Economics and Mathematical Systems*, 166, 1979.

[30] V. F. Demyanov and A. M. Rubinov. *Quasidifferentiability and related topics.* Springer, 2000.

[31] C. Derman. Denumerable state markovian decision processes average cost case. *Ann. Math. Statist.*, 37:1545–1554, 1966.

[32] U. Dini. *Fondamenti per la Teoria delle Funzioni di Variabili Reali.* Pisa, Italy, 1878.

[33] R. Dorfman, P. A. Samuelson, and R. M. Solow. *Linear programming and economic analysis.* Courier Corporation, 1958.

[34] A. Y. Dubovitskii and A. A. Milyutin. Extremum problems in the presence of restrictions, [in Russian]. *Zh. Vychisl. Mat. Mat. Fiz*, 5:1–80, 1965.

[35] H. Fast. Sur la convergence statistique. *Colloq. Math*, 2:241–244, 1951.

[36] A. Federgruen and H. C. Tijms. The optimality equation in average cost denumerable state semi-markov decision problems, recurrency conditions and algorithms. *J. Appl. Probab*, 15:356–373, 1978.

[37] W. Fenchel. Convex cones, sets and functions. 1951.

[38] V. Gaitsgory and S. Rossomakhine. Linear programming approach to deterministic long run average problems of optimal control. *SIAM journal on control and optimization*, 44, 2006.

[39] D. Gale. On optimal development in a multi-sector economy. *The Review of Economic Studies*, 34:1–18, 1976.

[40] R. N. Gasimov. Augmented lagranigian duality and nondifferentiable optimization methods in nonconvex programming. *Journal of global optimization*, 24:187–203, 2002.

[41] R. N. Gasimov and A. M. Rubinov. On augmented lagrangians for optimization problems with a single constraint. *Journal of global optimization*, 28:153–173, 2004.

[42] J. Gunawardena. An introduction to idempotency. *Basic Research Institute in the Mathematical Sciences HP Laboratories Bristol HPL-BRIMS-96-24*, 1996.

[43] J. Gunawardena. From max-plus algebra to non-expansive mappings: a non-linear theory for discrete event systems. *Theoretical Computer Science, Technical Report HPL-BRIMS-99-07, Hewlett-Packard Labs*, 1999.

[44] J. Gunawardena and M. Keane. On the existence of cycle times for some non-expansive maps. *Basic Research Institute in the Mathematical Sciences HP Laboratories Bristol HPL-BRIMS-95-03*, 1995.

[45] H. Halkin. Necessary conditions for optimal control problems with infinite horizons. *Econometrica*, 42:267–272, 1974.

[46] J. B. Hiriart-Urruty. Refinements of necessary optimality conditions in nondifferentiable programming i. *Applied mathematics and optimization*, 5:63–82, 1979.

[47] J. B. Hiriart-Urruty. Tangent cones, generalized gradients and mathematical programming in banach spaces. *Mathematics of operations research*, 4:79–97, 1979.

[48] J. B. Hiriart-Urruty. ϵ-subdifferential calculus, convex analysis and optimization. *(London), Pitman, Boston, Mass*, pages 43–92, 1982.

[49] W. J. Hopp, J. C. Bean, and R. L. Smith. A new optimality criterion for nonhomogeneous markov decision processes. *Operations Research*, 35:875–883, 1987.

[50] A. R. Howard. *Dynamic programming and Markov process.* Technology Press and Wiley, Nework, 1960.

[51] A. Ioffe and V. Tikhomirov. *Theory of extremal problems.* [in Russian], Nauka, Moskow, 1974.

[52] A. D. Ioffe. Regular points of lipschitz mappings. *Trans. Amer. Math. Soc*, 251:61–69, 1979.

[53] A. D. Ioffe. Nonsmooth analysis: Differential calculus of nondifferentiable mappings. *Trans. Amer. Math. Soc*, 266:1–56, 1981.

[54] A. D. Ioffe. Approximate subdifferentials and applications 1.the finite dimensional theory. *Trans. Amer. Math. Soc*, 281:289–316, 1984.

[55] A. D. Ioffe. Calculus of dini subdifferentials of functions and contingent coderivatives of set-valued maps. *Nonlinear Anal., Theory Methods Appl*, 8:517–539, 1984.

[56] A. D. Ioffe. Approximate subdifferentials and applications, III: The metric theory. *Mathematika*, 36:1–38, 1989.

[57] R. Kasimbeyli. A nonlinear cone separation theorem and scalarization in nonconvex vector optimization. *SIAM J. Optim*, 20:1591–1619, 2010.

[58] R. Kasimbeyli and M. Mammadov. On weak subdifferentials, directional derivatives, and radial epiderivatives for nonconvex functions. *SIAM Journal on Optimization*, 20:841–855, 2009.

[59] R. Kasimbeyli and M. Mammadov. Optimality conditions in nonconvex optimization via weak subdifferentials. *Nonlinear Anal.*, 74:2534–2547, 2011.

[60] A. Y. Kruger. Subdifferentials of nonconvex functions and generalized directional derivatives, [in Russian]. *Deposited at VINITI, Minsk*, pages 2661–77, 1977.

[61] A. Y. Kruger. ϵ-semidifferentials and ϵ-normal elements, [in Russian]. *Deposited at VINITI, Minsk*, pages 1331–81, 1981.

[62] A. Y. Kruger. Generalized differentials of nonsmooth functions and necessary conditions for an extremum. *Siberian Mathematical Journal*, 26(3):370–379, 1981.

[63] A. Y. Kruger. Generalized differentials of nonsmooth functions, [in Russian]. *Deposited at VINITI, Minsk*, pages 1332–81, 1981.

[64] A. Y. Kruger. Properties of generalized differentials. *Sib. Math. J.*, 26:822–832, 1985.

[65] A. Y. Kruger. On Fréchet subdifferentials. *Journal of Mathematical Sciences*, 116:3325–3358, 2003.

[66] A. Y. Kruger and B. S. Mordukhovich. Minimization of nonsmooth functionals in optimal control problems, [in Russian]. *Eng. Cybernetics*, 16:126–133, 1978.

[67] A. Y. Kruger and B. S. Mordukhovich. Extremal points and the euler equation in nonsmooth optimization, [in Russian]. *Dokl. Akad. Nauk BSSR*, 24:684–687, 1980.

[68] A. Y. Kruger and B. S. Mordukhovich. Generalized normals and derivatives and necessary optimality conditions in nondifferential programming, [in Russian]. *Deposited at VINITI*, 494-80, 1980.

[69] J. B. Lasserre. Decision horizon, overtaking and 1-optimality criteria in optimal control. *Advances in Optimization. Lecture Notes in Mathematics, no. 302*, 302:247–261, 1988.

[70] G. Lebourg. Valeur moyenne pour gradient généralisé. *C. R. Acad. Sér. Paris*, 281:795–797, 1975.

[71] C. Lee and E. Denardo. Rolling planning horizons: error bounds for the dynamic lot size model. *Mathematics of operation research*, 11:423–432, 1986.

[72] A. Leizarowitz. Infinite horizon autonomous systems with unbounded cost. *Appl Math Optim*, 13:19–43, 1985.

[73] D. T. Luc and S. Swaminathan. A characterization of convex functions. *Nonlinear Anal*, 20:697–701, 1993.

[74] V. Makarov and A. Rubinov. *Mathematical theory of economic dynamics and equilibria.* Springer-Verlag, New York, 1977.

[75] L. W. McKenzie. Turnpike theory. *Econometrica*, 44:841–866, 1976.

[76] H. Minkowski. Theorie der konvexen korper, insbesondere begrundung ihres ober flachenbegriffs. *Gesammelte Abhandlungen, II, Teubner*, 1911.

[77] G. J. Minty. Fonctionelles sous-différentiables. *C. R. Acad. Sci.*, 257:4117–4119, 1963.

[78] G. J. Minty. On the monotonicity of the gradient of a convex function. *Pacific J. Math*, 14:243–247, 1964.

[79] G. J. Minty. Proximité et dualité dans un espace hilbertien. *Bull. Soc. Math. France*, 93:273–299, 1965.

[80] H. Mohebi. Topical functions and their properties in a class of ordered banach spaces. pages 343–361, 2005.

[81] B. S. Mordukhovich. Maximum principle in problems of time optimal control with nonsmooth constraints. *J. Appl. Math. Mech*, 40:960–969, 1976.

[82] B. S. Mordukhovich. *Approximation Methods in Problems of Optimization and Control.* [in Russian], Nauka, Moscow, 1988.

[83] B. S. Mordukhovich. Variational analysis and generalized differentiation I: Basic theory, grundlehren der mathematischen wissenschaften. *Springer*, 2006.

[84] B. S. Mordukhovich and A. Y. Kruger. Necessary optimality conditions in a problem of terminal control with nonfunctional constraints, [in Russian]. *Dokl. Akad. Nauk BSSR*, 20:1064–1067, 1976.

[85] L. W. Neustadt. A general theory of extremals. *J. Comput. Syst. Sci*, 3:57–92, 1969.

[86] J. Penot. *Calculus Without Derivatives.* Graduate texts in mathematics. Springer, 2013.

[87] J. P. Penot. Calcul sousdifferentiel et optimisation. *J. Funct. Anal*, 27:248–276, 1978.

[88] B. N. Pshenichny. On necessary conditions for an extremum of nonsmooth functions, [in Russian]. *Kibernetika*, 6:92–96, 1977.

[89] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* Wiley, New York, 1994.

[90] S. Rayan. *Degenercy in discrete infinite horizon optimization.* PhD thesis, Department of industrial and operations engineering, The University of Michigan, Ann Arbor, 1988.

[91] S. M. Rayan, J. C. Bean, and R. L. Smith. A tie-breaking rule for discrete infinite horizon optimization. *Operation Research*, 40:117–126, 1992.

[92] R. T. Rockafellar. Characterization of subdifferentials of convex functions. *Pacific J. Math*, 17:497–510, 1966.

[93] R. T. Rockafellar. *Convex analysis.* Princeton University Press, New Jersey, 1970.

[94] R. T. Rockafellar and R. J. B. Wets. *Variational Analysis, Grundlehren Der Mathematischen Wissenschaften.* Springer, 1998.

[95] S. M. Ross. Non-discounted denumerable markovian decision models. *Ann. Math. Statist*, 39:412–2423, 1968.

[96] A. M. Rubinov and I. Singer. Topical and sub-topical functions, downward sets and abstract convexity. *Optim.*, 50:307–351, 2001.

[97] E. Sachs. Differentiability in optimization theory. *Math. Operationsforsch. Statist., Ser. Optimization*, 9:497–513, 1978.

[98] I. Schochetman and R. Smith. Optimality criteria for deterministic discrete-time infinite horizon optimization. *International Journal of Mathematics and Mathematical Sciences*, pages 57–80, 2004.

[99] I. Schochetman and R. Smith. Existence of efficient solutions in infinite horizon optimization under continuous and discrete controls. *Operations Research Letters*, 33:97–104, 2005.

[100] I. E. Schochetman and R. L. Smith. Infinite horizon optimization. *Mathematics of Operations Research*, 14:559–574, 1989.

[101] I. E. Schochetman and R. L. Smith. Finite dimensional approximation in infinite dimensional mathematical programming. *Mathematical Programming*, 54:307–333, 1992.

[102] I. E. Schochetman and R. L. Smith. Existence and discovery of average optimal solutions in deterministic infinite horizon optimization. *Mathematics of Operations Research*, pages 416–432, 1998.

[103] I. E. Schochetman and R. L. Smith. A finite algorithm for solving infinite dimensional optimization problems. *Annals of Operations Research*, 101:119–142, 2001.

[104] I. Singer. *Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces*. Springer, 1970.

[105] I. Singer. *Abstract convex analysis*. Wiley-Interscience, New York, 1997.

[106] I. Singer. Further applications of the additive min-type coupling function. *Optimization*, 51, 2002.

[107] I. Singer. Elementary topical functions on b-complete semimodules over b-complete idempotent semifields. *Lin. Alg. Appl.*, 433:2139–2146, 2010.

[108] I. Singer and V. Nitica. Topical functions on semimodules and generalizations. *Lin. Alg. Appl.*, 437:2471–2488, 2012.

[109] I. Singer and V. Nitica. Extended-valued topical and anti-topical functions on semimodules. *Lin. Alg. Appl.*, 446:25–70, 2014.

[110] J. V. Tiel. *Convex analysis: an introductory text*. John Wiley, 1984.

[111] H. C. Tijms. *A First Course in Stochastic Models*. Wiley, New York, 2003.

[112] A. F. Veinott. On finding optimal policies in discrete dynamic programming with no discounting. *Ann. Math. Statist.*, 37:1284–1294, 1966.

[113] A. O. Wachs, I. E. Schochetman, and R. L. Smith. Average optimality in non-homogeneous infinite horizon markov decision processes. *Mathematics of Operations Research*, 36:147–164, 2011.

[114] J. Warga. Controllability and necessary conditions in unilateral problems without differentiability assumptions. *SIAM J. Control Optim.*, 14:546–573, 1876.

[115] C. C. Weizsacker. Existence of optimal programs of accumulation for an infinite horizon. *The Review of Economic Studies*, 32:84–105, 1965.

[116] A. Zaslavski. *Turnpike properties in the calculus of variations and optimal control.* Springer, 2006.

[117] C. Zăalinescu. *Convex Analysis in General Vector Spaces.* World Scientific, 2002.